Stephan Trenn

Distributional Differential Algebraic Equations

Distributional Differential Algebraic Equations

Von Stephan Trenn

Universitätsverlag Ilmenau
2009

Impressum

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Angaben sind im Internet über http://dnb.d-nb.de abrufbar.

Diese Arbeit hat der Fakultät für Mathematik und Naturwissenschaften der Technischen Universität Ilmenau als Dissertation vorgelegen.

Tag der Einreichung:	29. Mai 2009
1. Gutachter:	Prof. Dr. Achim Ilchmann (Technische Universität Ilmenau)
2. Gutachter:	Prof. Dr. Hans Babovsky (Technische Universität Ilmenau)
3. Gutachter:	Prof. Dr. Daniel Liberzon (University of Illinois at Urbana-Champaign, USA)
Tag der Verteidigung:	28. Juli 2009

Technische Universität Ilmenau/Universitätsbibliothek
Universitätsverlag Ilmenau
Postfach 10 05 65
98684 Ilmenau
www.tu-ilmenau.de/universitaetsverlag

Herstellung und Auslieferung
Verlagshaus Monsenstein und Vannerdat OHG
Am Hawerkamp 31
48155 Münster
www.mv-verlag.de

ISBN 978-3-939473-57-2 (Druckausgabe)
urn:nbn:de:gbv:ilm1-2009000207

Abstract

Linear implicit differential equations of the form $E\dot{x} = Ax + f$ are studied. If the matrix E is not invertible, these equations contain differential as well as algebraic equations. Hence $E\dot{x} = Ax + f$ is called differential algebraic equation (DAE).

A main goal of this dissertation is the consideration of certain distributions (or generalized functions) as solutions and studying time-varying DAEs, whose coefficient matrices have jumps. Therefore, a suitable solution space is derived. This solution space allows to study the important class of switched DAEs.

The space of piecewise-smooth distributions is introduced as the solution space. For this space of distributions, it is possible to define a multiplication, hence DAEs can be studied whose coefficient matrices have also distributional entries. A distributional DAE is an equation of the form $E\dot{x} = Ax + f$ where the matrices E and A contain piecewise-smooth distributions as entries and the solutions x as well as the inhomogeneities f are also piecewise-smooth distributions.

For distributional DAEs, existence and uniqueness of solutions are studied, therefore, the concept of regularity for distributional DAEs is introduced. Necessary and sufficient conditions for existence and uniqueness of solutions are derived. As special cases, the equations $\dot{x} = Ax + f$ (distributional ODEs) and $N\dot{x} = x + f$ (pure distributional DAE) are studied and explicit solution formulae are given.

Switched DAEs are distributional DAEs with piecewise constant coefficient matrices. Sufficient conditions are given which ensure that all solutions of a switched DAE are impulse free. Furthermore, it is studied which conditions ensure that arbitrary switching between stable subsystems yield a stable overall system.

Finally, controllability and observability for distributional DAEs are studied. For this, it is accounted for the fact that input signals can contain impulses, hence an “instantaneous” control is theoretically possible. For a DAE of the form $N\dot{x} = x + bu$, $y = cx$, with constant, nilpotent N and constant vectors b and c, a normal form is given which allows for a simple characterization of controllability and observability.

Zusammenfassung

Lineare implizite Differentialgleichungen der Form $E\dot{x} = Ax + f$ werden untersucht. Da die Matrix E nicht als invertierbar angenommen wird, enthält das Gleichungssystem neben den Differentialgleichungen auch algebraische Gleichungen. Deshalb werden diese Gleichungen differential-algebraische Gleichungen (differential algebraic equations, DAEs) genannt.

Ein wesentliches Ziel der Dissertation ist es, Distributionen (oder verallgemeinerte Funktionen) als Lösungen zuzulassen und gleichzeitig soll es möglich sein, zeitvariante DAEs zu untersuchen, deren Koeffizientenmatrizen Sprünge haben können. Dazu wird zunächst ein geeigneter Lösungsraum hergeleitet. Insbesondere ist es mit diesem Lösungsraum möglich, die wichtige Klasse der geschalteten DAEs (switched DAEs) zu untersuchen.

Als Lösungsraum wird der Raum der stückweise glatten Distributionen (piecewise-smooth distributions) eingeführt. Für diesen Raum ist es möglich, eine Multiplikation zu definieren, so dass auch DAEs betrachtet werden können, deren Koeffizienten ebenfalls distributionelle Einträge haben. Eine distributionelle DAE ist eine Gleichung der Form $E\dot{x} = Ax + f$, bei der die Matrizen E und A stückweise glatte Distributionen als Einträge enthalten und die Lösungen x sowie die Inhomogenitäten f ebenfalls stückweise glatte Distributionen sind.

Für distributionelle DAEs wird die Existenz und Eindeutigkeit von Lösungen untersucht, dazu wird das Konzept der Regularität für distributionelle DAEs eingeführt. Es werden notwendige und hinreichende Bedingungen für die Existenz und Eindeutigkeit von Lösungen hergeleitet. Als Spezialfälle werden die beiden Gleichungen $\dot{x} = Ax + f$ (so genannte distributionelle ODEs) und $N\dot{x} = x + f$ (so genannte reine distributionelle DAEs) untersucht, für die explizite Lösungsformeln angegeben werden können.

Geschaltete DAEs sind distributionelle DAEs mit stückweise konstanten Koeffizientenmatrizen. Es werden hinreichende Bedingung hergeleitet, die sicherstellen, dass die Lösungen von geschalteten DAEs keine Impulse enthalten. Weiterhin wird untersucht, unter welchen Bedingungen das beliebige Schalten zwischen stabilen Teilsystemen zu

einem stabilen Gesamtsystem führt.

Schließlich werden Steuerbarkeit und Beobachtbarkeit für distributionelle DAEs untersucht. Hierbei wird berücksichtigt, dass das Eingangssignal Impulse enthalten kann und damit theoretisch eine „instantane“ Steuerung möglich ist. Für eine DAE der Form $N\dot{x} = x + bu$, $y = cx$, mit konstanten, nilpotenten N sowie konstanten Vektoren b und c wird eine Normalform angegeben, die eine einfache Charakterisierung der Steuerbarkeit und Beobachtbarkeit ermöglicht.

Danksagung

Ohne Achim würde es diese Arbeit nicht geben, deshalb gilt vor allem ihm mein ausdrücklicher Dank für alles, insbesondere dass er mich vom ersten Semester an gefordert und gefördet hat. Er war in den letzten neun Jahren nicht nur ein ausgezeichneter wissenschaftlicher Mentor, sondern er ist auch ein sehr guter persönlicher Freund geworden.

Bei Aylin möchte ich mich bedanken, weil sie mir in den letzten Jahren in stressigen Situationen den Rücken frei gehalten hat und weil sie natürlich einfach eine tolle Frau ist, die mir auch außerhalb der mathematischen Welt Freude bereitet.

Zum Ende wurde es zeitlich doch recht eng, deshalb bedanke ich mich bei den Gutachtern (Prof. Dr. Hans Babovsky, Prof. Dr. Daniel Liberzon und natürlich Achim), dass sie zügig die Gutachten erstellt haben. Natürlich bedanke ich mich auch dafür, dass sie sich überhaupt die Zeit genommen haben, meine Arbeit detailiert zu lesen und nachzuvollziehen.

Dank gilt auch allen, mit denen ich in den letzten Jahren wissenschaftlich diskutiert habe, was (mehr oder weniger) zur endgültigen Fassung dieser Dissertation beigetragen hat. Hervorzuheben sind hier vor allem Prof. Dr. Jan C. Willems (K.U. Leuven), Prof. Dr. Volker Mehrmann und Dr. Timo Reis (beide von der TU Berlin) sowie, aus Ilmenau, Prof. Dr. Armin Hoffmann, Prof. Dr. Carsten Trunk, PD Dr. Hans Crauel (jetzt in Frankfurt/Main) und B.Sc. Thomas Berger. Dafür, dass (neben den interessanten wissenschaftlichen Diskussionen) die Büroarbeit nicht zu langweilig und eintönig wurde, bedanke ich mich bei meinen Zimmerkollegen Markus und Norman.

Mit Abschluss der Promotion endet auch meine „Ilmenauer Zeit“, die rückblickend betrachtet eine sehr schöne Zeit war. Deshalb bedanke ich mich zunächst bei Ilmenau als Universitätsstadt ganz allgemein. Speziell möchte ich mich bei den „Gremienstudenten“ aber auch „Gremienprofessoren“ bedanken, denn die Gremienarbeit hat einerseits meine geistigen Horizont deutlich erweitert und es hat außerdem viel Spaß gemacht, in den Tiefen der Universitätsstrukturen zu wühlen, auch wenn man nicht immer das erreicht hat, was man wollte. Speziell hervorheben möchte ich auch das Eiscafe Venezia, weil es kreative Pausen mit

Kaffee oder Eis ermöglichte, die natürlich unerlässlich für produktives wissenschaftliches Arbeiten sind.

Schließlich bedanke ich mich noch bei meinen Eltern, die zwar nicht direkt zur Entstehung der Dissertation beigetragen haben, aber immer für mich da waren, wenn ich sie gebraucht habe.

Contents

1 Introduction

In this dissertation linear implicit differential equations or *differential algebraic equations (DAEs)* of the form

$$E\dot{x} = Ax + f$$

are studied. In the simplest case, $E, A \in \mathbb{R}^{m\times n}$ are constant matrices, $f : \mathbb{R} \to \mathbb{R}^m$ is some inhomogeneity and solutions are differentiable functions $x : \mathbb{R} \to \mathbb{R}^n$. It is not assumed that E is invertible which yields that, in addition to differential equation, also algebraic equations are involved. DAEs naturally occur when modelling linear electrical circuits, simple mechanical systems or, in general, (linear) systems with additional (linear) algebraic constraints.

On a first view it seems that the theory of differential algebraic equations is well developed and mature (see e.g. the textbooks [Cam80, Cam82, GM86, Dai89, Apl91, RR02, KM06] and the references therein). However, a solution theory for switched DAEs or, in general, for DAEs with time-varying *discontinuous* coefficient matrices seems not to be available. One aim of this dissertation is therefore to develop a solution framework for time-varying DAEs with discontinuous coefficients and to establish a starting point for future research on, e.g., switched DAEs.

1.1 Distributional solutions

If only classical solutions (i.e. differential functions) of DAEs are considered, then many interesting and important properties of DAEs (e.g. inconsistent initial values, impulsive solutions, impulse-controllability and -observability) can not be studied even in the constant coefficient case. There are basically two well known approaches too enlarge the solution space: 1) Considering solutions where only certain components are differentiable, this leads to so called properly stated leading term introduced in [BM02], and 2) the possible solution space for DAEs is enlarged to allow for so called *generalized functions* or *distributions* [VLK81, Cam82]. The latter approach has the advantage that the underlying solution space does not depend on the considered DAE and is therefore conceptionally easier. Furthermore, the first approach alone cannot be used to study impulse-related questions.

The disadvantage of a distributional solution space is that distributions are not functions anymore, so that for example an evaluation at a certain time is in general not possible, in particular, initial value problems cannot be formulated directly. In fact, the space of distributions is for practical purposes too large, in most situations it is enough that the Dirac impulse (δ-function) and its derivatives are allowed as solutions. Therefore, Cobb [Cob84] introduced the smaller space of *piecewise-continuous distributions*, for which it is possible to define a *restriction* of a distribution which he used to formulate initial value problems and to define the impulsive part of a distribution. However, the derivatives of a piecewise-continuous distributions are no longer piecewise-continuous so that the problems of a distributional approach are only solved partially. Another approach is the solution space of *impulsive-smooth distributions* as introduced in [RR96b] which had its origins in [HS83], this space consists of distributions which are smooth functions on $\mathbb{R} \setminus \{0\}$ and which can have Dirac impulses and its derivatives at zero. In this setup initial value problems are firstly reformulated as initial trajectory problems and, secondly, the inhomogeneity is changed such that the initial trajectory becomes consistent. A disadvantage of the space of impulsive-smooth distributions is that Dirac-impulses can only occur at zero, because a general distributional solution theory should allow for arbitrarily many Dirac-impulses at arbitrary times.

There seems to be no literature on DAEs with non-continuous coefficients in combination with distributional solutions. In fact, this is not surprising because for distributions only a multiplication with smooth functions is well defined. However, switched DAEs can also be written as DAEs which should only be valid on certain intervals, which bypasses the need to multiply with discontinuous functions. Initial value problems (in particular with inconsistent initial values) can also be seen as a DAEs which should be valid only on the interval $[t_0, \infty)$.

This motivates studying restrictions of distributions. Surprisingly, defining a restriction for distributions turns out to be difficult, in fact, it can be shown that it is not possible to define a restriction for distributions in general (Theorem 2.2.2). Therefore, the space of piecewise-regular distributions is defined on which a restriction can be defined. As a consequence, it is then also possible to define the multiplication with a certain class of non-continuous functions. However, it is still not possible to study switched DAEs because, as with the approach in [Cob84], the derivative of an arbitrary piecewise-regular distribution is not always piecewise-regular.

This leads to the definition of the space of *piecewise-smooth distributions* which combines the advantages of the space of piecewise-continuous distributions (Dirac impulses can be everywhere) and the space of impulsive-smooth distributions (closed under differentiation).

The piecewise-smooth distributional framework can then be applied to switched DAEs (i.e. DAEs with piecewise-constant coefficient matrices), see Section 4 where, as an application of the distributional solution theory, conditions are formulated which ensure that the solutions of the switched DAEs do not exhibit impulsive behaviour. To illustrate the relevance of the proposed framework a "real world" electrical circuit is studied in detail with respect to its ability to generate impulses in response to switches or component failures. Another application of the proposed framework is the stability analysis of switched DAEs.

1.2 Distributional DAEs and multiplication of distributions

For the analysis of classical DAEs (i.e. with constant coefficients) equivalence transformations play an important role. For example, multiplying a DAE from the left with an invertible matrix does not change its solutions but it might reveal special structural properties of the DAE, the same is true for a coordinate transformation. Hence many properties of the classical DAE $E\dot{x} = Ax + f$ can also be obtained by studying the "equivalent" DAE $SET\dot{x} = SATx + Sf$, where S, T are invertible matrices. An example for this is the Kronecker normal form [Kro90, Gan59] or, for regular matrix pairs, the Weierstraß normal form [Wei68, Gan59].

If time-varying coefficient matrices are considered it is natural to also consider time-varying transformation matrices, hence the matrix pair (E, A) is transformed to the matrix pair $(SET, SAT - SET')$, where T' is the time-derivative of the time-varying coordinate transformation. If T is piecewise-smooth (as the matrices E and A) then T' will only be well defined in a distributional sense. Hence if all transformation matrices S, T are allowed which have the same type as the coefficient matrices (i.e. matrices of piecewise-smooth functions) then the occurrence of T' implies that in the coefficient matrices of DAEs also Dirac impulses must be allowed.

This leads to the problem that a multiplication of the Dirac impulse with a (piecewise-smooth) distribution must be defined. However, it is well known that it is not possible to define a multiplication of distributions in general and even for the simple product of the Dirac impulse with itself there has been a considerable dispute in the literature, whether the square of the Dirac impulse is well defined or not (see Remark 2.4.5). Furthermore, allowing Dirac impulses in the coefficient matrices of a DAE implies inductively with the same argument as above that all derivatives of the Dirac impulse must be allowed in the coefficient matrices as well.

In summary: if one wants to study "natural" transformation of time-varying DAEs with piecewise-smooth entries, then it is necessary to enlarge the system class to encompass also coefficient matrices whose entries are piecewise-smooth distributions and to define a multiplication *for piecewise-smooth distribution.*

In fact, defining a suitable multiplication is possible, although there are two ways to define the multiplication, one way yields "causal" DAEs, the other one yields "anticausal" DAEs. The causal multiplication is called *Fuchssteiner multiplication* because Fuchssteiner studied a very similar multiplication for distributions [Fuc68, Fuc84].

Therefore, the class of *distributional DAEs* $E\dot{x} = Ax + f$, where the coefficient matrix entries, the inhomogeneities and the solutions are piecewise-smooth distributions are well defined and can be studied.

The existence of Dirac impulses in the coefficient matrices can also be motivated by so-called impulsive systems (see e.g. [LBS89]) which can be rewritten in closed form as $\dot{x} = Ax$, where A has Dirac impulses at the jump-times of the state.

1.3 Regularity of distributional DAEs

As a first application of the piecewise-smooth distributional framework, the well known concept of regularity for classical DAEs is generalized to to distributional DAEs in Section 3. Roughly speaking, a distributional DAE is called regular if, and only if, existence and uniqueness of solutions are guaranteed for arbitrary initial values and inhomogeneities. Sufficient and necessary conditions for regularity are given. For two special DAEs, so called distributional ODEs and pure distributional DAEs, explicit solution formulae are developed. Furthermore, a generalized Weierstraß normal form is proposed.

A direct consequence is that the regularity of time-varying DAEs with analytical coefficient matrices in the sense of *analytical solvability* [CP83] implies regularity of the corresponding distributional DAE (Corollary 3.5.4).

The regularity for distributional DAEs implies, by Theorem 3.2.5, regularity for time-varying DAEs as in [RR96a, Defn. 3.1], but it is not clear how regularity relates to the concept of *complete regularity* [RR96a, Defn. 3.4]. Furthermore, (complete) regularity in the sense of [RR96a] does not imply uniqueness of solutions, see Example 3.2.7.

1.4 Controllability and observability

For DAEs the controllability definition split into two independent definitions: R-controllability and impulse-controllability; the same is true for the definition of observability (see e.g. [Dai89]). These definitions are somewhat mysterious (in particular the impulse-controllability and -observability definitions), because no proper distributional framework is proposed. In Section 5 jump-controllability, impulse-controllability, jump-observability and impulse-controllability are defined in such a way that 1) the definition incorporate the time-varying nature of the DAE, 2) jump- and impulse-controllability as well as jump- and impulse-observability are defined such that they are in some sense complementary, 3) for classical DAEs the definitions are equivalent to the classical definitions.

Finally, for a pure DAE with constant coefficients a new normal form is proposed which separates the state into four substates which are impulse-controllable and impulse-observable, impulse-observable but not impulse-controllable, impulse-controllable but not impulse-observable, neither impulse-controllable nor impulse-observable. The normal form is also used to construct a control input such that an impulse-controllable DAE does not have an impulsive solution, for an impulse-observable DAE it is shown how the impulses in the solution can be determined by the output.

1.5 Previously published results and joint work

The following parts of this dissertation are already published or submitted for publication. Parts of Sections 3.3.1 and 3.4 are published in [Tre08a] (without proofs). The submitted manuscript [Tre08b] contains parts of Sections 2, 3.1, 3.2 and 3.5. The results of Sections 4.2.2 and 4.2.3 are submitted for publication [Tre09a]. The normal from from Section 5.3 is published in [Tre09b].

The Quasi-Weierstraß form in Section 4.2.1 stems from a joined work with Thomas Berger and Achim Ilchmann (both Ilmenau University of Technology) which is submitted for publication [BIT09]. The stability results for switched systems in Section 4.3 were obtained in cooperation with Daniel Liberzon (University of Illinois at Urbana-Champaign) and are accepted for publication [LT09]. The example of a dual redundant buck converter in Section 4.2.4 was provided by Alejandro D. Domínguez-García (University of Illinois at Urbana-Champaign).

1.6 Basic notational conventions

The real numbers, complex number, natural numbers and integers are denoted by $\mathbb{R}, \mathbb{C}, \mathbb{N}, \mathbb{Z}$, respectively. Positive numbers and non-negative numbers are denoted by the indices $_{>0}$ or $_{\geq 0}$, in particular $\mathbb{Z}_{\geq 0} = \mathbb{N}$.

The subset-relation is denoted by $\subseteq$ while the proper subset-relation is $\subset$. For some set $M \subseteq \mathbb{R}$ the indicator function $\mathbb{1}_M : \mathbb{R} \to \{0, 1\}$ is given by $\mathbb{1}_M(t) = 1$ if, and only if, $t \in M$.

For some vectors $v_1, v_2, \ldots, v_m \in \mathbb{R}^n$ denote the matrix consisting of this vectors as columns as $[v_1, v_2, \ldots, v_m] \in \mathbb{R}^{n\times m}$, the matrix consisting of these vectors as rows is denoted by $[v_1/v_2/\cdots/v_m] \in \mathbb{R}^{m\times n}$. A matrix $N \in \mathbb{R}^{n\times n}$ is called nilpotent if, and only if, $N^\nu = 0$ for some $\nu \in \mathbb{N}$.

For some matrix $M \in \mathbb{R}^{m\times n}$ and set $\mathcal{M} \subseteq \mathbb{R}^n$ the image of $\mathcal{M}$ under M is $M\mathcal{M} := \{\ Mx \mid x \in \mathcal{M}\ \}$ and for $\mathcal{M} \subseteq \mathbb{R}^m$ the preimage of $\mathcal{M}$ under M is $M^{-1}\mathcal{M} := \{\ x \in \mathbb{R}^n \mid Mx \in \mathcal{M}\ \}$. In particular, $\ker M := M^{-1}\{0\}$ denotes the kernel of M and $\operatorname{im} M := M\mathbb{R}^n$ is the image of M. The direct sum of two linear subspaces is denoted by $\oplus$.

It is assumed, that the real numbers are equipped with the Lebesgue measure and that integrals are Lebesgue integrals.

2 Distribution theory

2.1 Review of classical distribution theory

Definition 2.1.1 (Test functions)
Let

$$\begin{aligned} \mathcal{C}^\infty &:= \mathcal{C}^\infty(\mathbb{R} \to \mathbb{R}) \\ &:= \{\ f : \mathbb{R} \to \mathbb{R} \mid f \text{ is arbitrarily often differentiable } \} \end{aligned}$$

be the space of *smooth functions*. The *support* of $f \in \mathcal{C}^\infty$ is given by

$$\operatorname{supp} f := \operatorname{cl} \{\ x \in \mathbb{R} \mid f(x) \neq 0\ \},$$

where $\operatorname{cl} M$ is the closure of the set $M \in \mathbb{R}$. The space of *test functions* is defined by

$$\mathcal{C}_0^\infty := \mathcal{C}_0^\infty(\mathbb{R} \to \mathbb{R}) := \{\ \varphi \in \mathcal{C}^\infty \mid \operatorname{supp} \varphi \text{ is bounded } \},$$

i.e. $\mathcal{C}_0^\infty$ is the space of smooth functions with bounded support. □

It can be shown that the space of test functions $\mathcal{C}_0^\infty$ is a topological space (see e.g. [Jan71, §12] or [Wer02, VIII.1.Bsp.(f)]). In the following it is assumed that $\mathcal{C}_0^\infty$ is equipped with the topology given in [Jan71, §12].

Lemma 2.1.2 ([Jan71, Sätze 12.7 and 14.2])
A linear operator $L : \mathcal{C}_0^\infty \to \mathbb{R}$ is continuous if, and only if, $L(\varphi_n) \to 0$ as $n \to \infty$ for all sequences $(\varphi_n)_{n\in\mathbb{N}} \in (\mathcal{C}_0^\infty)^\mathbb{N}$ with

(i) $\exists$ compact $K \subseteq \mathbb{R}\ \forall n \in \mathbb{N} : \operatorname{supp} \varphi_n \subseteq K$, and

(ii) $\forall i \in \mathbb{N} : \varphi_n^{(i)} \to 0$ uniformly as $n \to \infty$. □

Definition 2.1.3 (Distributions)
The space of *distributions* is given by

$$\mathbb{D} := \{\ D : \mathcal{C}_0^\infty \to \mathbb{R} \mid D \text{ is linear and continuous } \}.$$ □

Note that the space of test functions is often denoted by $\mathcal{D}$ and the space of distribution is then defined as the dual space of $\mathcal{D}$ and is therefore denoted by $\mathcal{D}'$. However, nearly all statements in this work are made on distributions and the space of distributions plays a much more important role than the space of test functions. In particular, many times the space of distribution will have sub- and super-indices and it will improve readability to have a single letter for the space of distributions, therefore instead of $\mathcal{D}'$ the notation $\mathbb{D}$ is used.

Definition 2.1.4 (Regular distributions)
Let

$$L_{1,\text{loc}} := L_{1,\text{loc}}(\mathbb{R} \to \mathbb{R}) := \{\ f : \mathbb{R} \to \mathbb{R} \mid f \text{ is locally integrable }\}$$

be the space of *locally integrable functions*, i.e. the space of all measurable functions $f : \mathbb{R} \to \mathbb{R}$ for which the integral $\int_K |f|$ is finite for all compact sets $K \subseteq \mathbb{R}$. The *regular distribution* induced by $f \in L_{1,\text{loc}}$ is

$$f_{\mathbb{D}} : \mathcal{C}_0^\infty \to \mathbb{R}, \quad \varphi \mapsto \int_{\mathbb{R}} \varphi f.$$

The space of regular distributions is given by

$$\mathbb{D}_{\text{reg}} := \{\ f_{\mathbb{D}} \mid f \in L_{1,\text{loc}}\ \}. \qquad \square$$

Proposition 2.1.5 ([Jan71, Sätze 15.1 and 15.3])

$$\mathbb{D}_{\text{reg}} \subset \mathbb{D}$$

and

$$\forall\, f, g \in L_{1,\text{loc}} : \quad [\ f_{\mathbb{D}} = g_{\mathbb{D}} \quad \Leftrightarrow \quad f = g \text{ almost everywhere }]. \qquad \square$$

Definition 2.1.6 (Distributional derivative)
For $D \in \mathbb{D}$ let

$$\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} : \mathbb{D} \to \mathbb{D}, D \mapsto \big(\varphi \mapsto -D(\varphi')\big)$$

be the *distributional derivative* of D. The following notations for the derivative of D are also used:

$$D' := \dot{D} := \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} D. \qquad \square$$

Proposition 2.1.7 ([Jan71, Satz 19.1])
The distributional derivative is well defined, i.e. $D' \in \mathbb{D}$ for all distributions $D \in \mathbb{D}$, and it is a generalization of the classical derivative, i.e. for all differentiable functions $f : \mathbb{R} \to \mathbb{R}$ it holds that

$$(f')_{\mathbb{D}} = (f_{\mathbb{D}})'. \qquad \square$$

A consequence of this proposition is that distributions can arbitrarily often be differentiated, for higher derivatives the following notation is used:

$$(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{n+1} D := D^{(n+1)} := (D^{(n)})', \text{ where } D^{(0)} := D,\ n \in \mathbb{N} \text{ and } D \in \mathbb{D}.$$

Proposition 2.1.8 (Sequences of distributions, [Jan71, 28.3, 28.1])
Let $(D_n) \in \mathbb{D}^{\mathbb{N}}$ be a sequence of distributions such that for all $\varphi \in \mathcal{C}_0^\infty$ the sequence $\big(D_n(\varphi)\big) \in \mathbb{R}^{\mathbb{N}}$ converges. Then $D := \varphi \mapsto \lim_{n\to\infty} D_n(\varphi)$ is a distribution and (D_n) *converges* to D in the sense:

$$D_n \to D \text{ as } n \to \infty \quad :\Leftrightarrow \quad \forall \varphi \in \mathcal{C}_0^\infty : \lim_{n\to\infty} D_n(\varphi) = D(\varphi).$$

Furthermore,

$$D_n \to D \text{ as } n \to \infty \quad \Rightarrow \quad D_n' \to D' \text{ as } n \to \infty. \qquad \square$$

From a functional analysis viewpoint the above convergence is the well known weak* convergence.

Definition 2.1.9 (Multiplication with smooth functions)
Let $\alpha \in \mathcal{C}^\infty$ and $D \in \mathbb{D}$, then the multiplication αD is defined by

$$\alpha D : \mathcal{C}_0^\infty \to \mathbb{R},\ \varphi \mapsto D(\alpha\varphi). \qquad \square$$

Proposition 2.1.10 (Multiplication with smooth functions)
Let $\alpha \in \mathcal{C}^\infty$ and $D \in \mathbb{D}$, then

$$\alpha D \in \mathbb{D}$$

and the multiplication is a generalization of the standard multiplication, i.e.

$$\alpha f_{\mathbb{D}} = (\alpha f)_{\mathbb{D}} \quad \text{for all } f \in L_{1,\text{loc}}.$$

Furthermore,

$$(\alpha D)' = \alpha' D + \alpha D' \qquad \square$$

Proof. The first two assertions are shown in [Jan71, Satz 18.1]. The product rule of the differentiation follows from the definition:

$$\begin{aligned} \forall \varphi \in \mathcal{C}_0^\infty : (\alpha D)'(\varphi) &= -(\alpha D)(\varphi') = D(-\alpha\varphi') = D\big(\alpha'\varphi - (\alpha\varphi)'\big) \\ &= D(\alpha'\varphi) + D'(\alpha\varphi) = \alpha' D(\varphi) + \alpha D'(\varphi) \end{aligned} \qquad \boxed{\text{qed}}$$

Definition 2.1.11 (Support of a distribution)
The *support* of a distribution $D \in \mathbb{D}$ is defined by

$$\operatorname{supp} D := \mathbb{R} \Big\backslash \bigcup \left\{ O \subseteq \mathbb{R} \;\middle|\; \begin{array}{l} O \text{ open and } \forall \varphi \in \mathcal{C}_0^\infty : \\ \operatorname{supp} \varphi \subseteq O \;\Rightarrow\; D(\varphi) = 0 \end{array} \right\},$$

i.e. the support is the complement of the largest open set on which D vanishes. The set of all distributions with support in some $M \subseteq \mathbb{R}$ is

$$\mathbb{D}_M := \{\, D \in \mathbb{D} \mid \operatorname{supp} D \subseteq M \,\}. \qquad \square$$

Proposition 2.1.12 (Properties of Dirac impulses, [Jan71, §15])
The *Dirac impulse* at $t \in \mathbb{R}$ given by

$$\delta_t : \mathcal{C}_0^\infty \to \mathbb{R}, \; \varphi \mapsto \delta_t(\varphi) := \varphi(t)$$

is a distribution which is not regular. The support of δ_t and of all its derivatives is $\{t\}$, i.e.

$$\forall n \in \mathbb{N}\ \forall t \in \mathbb{R}: \ \delta_t^{(n)} \in \mathbb{D}_{\{t\}}.$$

The Dirac impulse is the distributional derivative of the so called *Heaviside function* $\mathbb{1}_{[0,\infty)}$, i.e.

$$\forall t \in \mathbb{R}: \ \delta_t = \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}(\mathbb{1}_{[t,\infty)})_{\mathbb{D}}.$$

For every distribution D with *point support*, i.e. $\exists t \in \mathbb{R} : \operatorname{supp} D \subseteq \{t\}$, there exists $N \in \mathbb{N}$, $a_0, a_1, \ldots, a_N \in \mathbb{R}$ such that

$$D = \sum_{i=0}^{N} a_i \delta_t^{(i)}.$$

□

Remark 2.1.13 (Product of smooth functions with Dirac impulses)
For $\alpha \in \mathcal{C}^\infty$ and $t \in \mathbb{R}$ it follows inductively from Definition 2.1.6 and Proposition 2.1.10 that for all $n \in \mathbb{N}$

$$\alpha\delta_t^{(n)} = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \alpha^{(i)}(t)\delta_t^{(n-i)}.$$

In particular, $\alpha\delta_t^{(n)} \in \mathbb{D}_{\{t\}}$. □

Remark 2.1.14 (Linear independence of Dirac impulses)
The elements of the set $\left\{ \delta_t^{(i)} \;\middle|\; t \in \mathbb{R}, i \in \mathbb{N} \right\}$ are linearly independent in the $\mathbb{R}$ vector space $\mathbb{D}$. In particular, for a fixed $t \in \mathbb{R}$ and for all $a_1, \ldots, a_n \in \mathbb{R}$, $n \in \mathbb{N}$,

$$\sum_{i=0}^{N} a_i \delta_t^{(i)} = 0 \quad \Leftrightarrow \quad a_1 = a_2 = \ldots = a_n = 0,$$

and for every finite or infinite set of distributions with pairwise disjoint point support, only the trivial linear combination is the zero distribution. □

Proposition 2.1.15 (Support of regular distributions)

$$\forall\, f \in L_{1,\mathrm{loc}}(\mathbb{R} \to \mathbb{R}): \quad [\, \operatorname{supp} f_{\mathbb{D}} \text{ has measure zero} \quad \Leftrightarrow \quad f_{\mathbb{D}} = 0 \,],$$

i.e. every non-trivial regular distribution has an essential support. □

Proof. Let $S = \operatorname{supp} f_{\mathbb{D}}$, then by definition $T = \mathbb{R}\backslash S$ is an open set and for all $\varphi \in \mathcal{C}_0^\infty$ with $\operatorname{supp}\varphi \subseteq T$ it is $f_{\mathbb{D}}(\varphi) = 0$. Every open subset of $\mathbb{R}$ can be written as a countable disjoint union of open intervals, hence there exists $l_i, r_i \in \mathbb{R}$, $i \in \mathbb{N}$, such that

$$T = \dot{\bigcup_{i\in\mathbb{N}}} (l_i, r_i).$$

Note that one runs into notational difficulties if T contains one or two unbounded intervals. This can be fixed by just considering $S \cup \mathbb{Z}$ instead of S.
Let $\varphi \in \mathcal{C}_0^\infty$; it will be shown, that $f_{\mathbb{D}}(\varphi) = 0$.
For $a, b \in \mathbb{R}$ and $\varrho > 0$ let $\mathbb{1}^{\varrho}_{[a,b]} \in \mathcal{C}_0^\infty$ be such that

$$\mathbb{1}^{\varrho}_{[a,b]}(t) = \begin{cases} 1, & t \in [a,b], \\ 0, & t \notin (a-\varrho, b+\varrho), \end{cases} \tag{2.1.1}$$

and $0 \le \mathbb{1}^{\varrho}_{[a,b]}(t) \le 1$ for all $t \in \mathbb{R}$ (for details how to construct such a function see e.g. [KR95, Satz 1.4]). For $\varepsilon > 0$ choose $\varepsilon_i > 0$, $i \in \mathbb{N}$, such that

$$\varepsilon_i < \min\left\{ \frac{\varepsilon}{2^{i+2}}, \frac{r_i - l_i}{4} \right\}.$$

Let

$$\varphi_\varepsilon := \varphi \prod_{i\in\mathbb{N}} \left(\mathbb{1} - \mathbb{1}^{\varepsilon_i}_{[l_i+\varepsilon_i, r_i-\varepsilon_i]} \right),$$

then it follows that $\varphi_\varepsilon \in \mathcal{C}_0^\infty$ and

$$\operatorname{supp}(\varphi - \varphi_\varepsilon) \subseteq T.$$

Hence $f_\mathbb{D}(\varphi - \varphi_\varepsilon) = 0$, or, equivalently, $f_\mathbb{D}(\varphi) = f_\mathbb{D}(\varphi_\varepsilon)$. Note further that

$$\operatorname{supp} \varphi_\varepsilon \subseteq S \,\dot\cup\, T_\varepsilon,$$

where

$$T_\varepsilon = \bigcup_{i\in\mathbb{N}} (l_i, l_i + 2\varepsilon_i) \cup (r_i - 2\varepsilon_i, r_i).$$

This, together with $\left|\varphi_\varepsilon(t)\right| \leq \left|\varphi(t)\right|$, yields

$$\begin{aligned}\left|f_\mathbb{D}(\varphi)\right| = \left|f_\mathbb{D}(\varphi_\varepsilon)\right| \leq \int_\mathbb{R} |\varphi_\varepsilon||f| = \int_S |\varphi_\varepsilon||f| + \int_{T_\varepsilon} |\varphi_\varepsilon||f| + \int_{T\setminus T_\varepsilon} |\varphi_\varepsilon||f| \\ \overset{(*)}{=} \int_{T_\varepsilon} |\varphi_\varepsilon||f| \leq \|\varphi\|_\infty \int_{T_\varepsilon} |f|,\end{aligned}$$

where the equality $(*)$ follows from the two facts that S has measure zero and that φ_ε is zero on $T\setminus T_\varepsilon$. The Lebesgue measure $\lambda(T_\varepsilon)$ of T_ε is

$$\lambda(T_\varepsilon) = \sum_{i\in\mathbb{N}} 4\varepsilon_i < \sum_{i\in\mathbb{N}} \frac{\varepsilon}{2^i} = \varepsilon,$$

hence $\int_{T_\varepsilon} |f|$ tends to zero if ε tends to zero. Therefore, $f_\mathbb{D}(\varphi) = 0$ and since φ was arbitrarily chosen it follows that $f_\mathbb{D} = 0$. qed

Proposition 2.1.16 ([KR95, Folg. 3.24])
Let $D \in \mathbb{D}_M$ for some $M \subseteq \mathbb{R}$ and let $\varphi \in \mathcal{C}_0^\infty$ with $\varphi^{(i)}(t) = 0$ for all $t \in M$ and all $i \in \mathbb{N}$. Then $D(\varphi) = 0$. □

Note that in Proposition 2.1.16 it is not assumed that $\operatorname{supp} \varphi \cap \operatorname{supp} D = \emptyset$.

Corollary 2.1.17 (Zero product)
Let $D \in D_M$ for some measurable $M \subseteq \mathbb{R}$ and let $\alpha \in \mathcal{C}^\infty$ with $\alpha^{(i)}(t) = 0$ for all $t \in M$ and all $i \in \mathbb{N}$. Then $\alpha D = 0$. □

Proof. For all $\varphi \in \mathcal{C}_0^\infty$ it is $(\alpha\varphi)^{(i)}(t) = 0$ for all $t \in M$ and all $i \in \mathbb{N}$. Hence, by Proposition 2.1.16, $(\alpha D)(\varphi) = D(\alpha\varphi) = 0$ qed

Proposition 2.1.18 (Distributional antiderivative, [Jan71, 22.4])
For every distribution $D \in \mathbb{D}$ there exists a *distributional antiderivative* $H \in \mathbb{D}$ of D, i.e. $H' = D$. For two distributional antiderivatives $H_1, H_2 \in \mathbb{D}$ of D there exists a constant $c \in \mathbb{R}$, such that $H_1 - H_2 = c\mathbb{1}_{\mathbb{D}}$, i.e. a distributional antiderivative is unique modulo a constant. □

2.2 Piecewise-regular distributions

2.2.1 Restrictions of distributions

The aim of this subsection is to define a *distributional restriction* in the following sense. In general, a distributional restriction is a mapping

$$\{\ M \subseteq \mathbb{R} \mid M \text{ measurable }\} \times \mathbb{D} \to \mathbb{D}, \quad (M, D) \mapsto D_M, \qquad (2.2.1)$$

i.e. the restriction should be defined again on the whole space of test function (and not only the subspace of test functions whose support is contained in $M \subseteq \mathbb{R}$). Furthermore, the distributional restriction should fulfill the following properties:

(R1) The distributional restriction (2.2.1) fulfills $D_M \subseteq \mathbb{D}_{\operatorname{cl} M}$ for all $D \in \mathbb{D}$, $M \subseteq \mathbb{R}$ and is for each fixed $M \subseteq \mathbb{R}$ a projection, i.e. $D \mapsto D_M$ is linear and idempotent.

(R2) For $f \in L_{1,\text{loc}}$ and measurable $M \subseteq \mathbb{R}$ let $f_M := \mathbb{1}_M f$, then the distributional restriction (2.2.1) fulfills

$$(f_M)_{\mathbb{D}} = (f_{\mathbb{D}})_M,$$

i.e. it is a generalization of restrictions of functions.

(R3) The restriction property of (2.2.1) for trivial cases is fulfilled, i.e. for all test functions $\varphi \in \mathcal{C}_0^\infty$, distributions $D \in \mathbb{D}$ and measurable sets $M \subseteq \mathbb{R}$ the following two implications hold

$$\operatorname{supp} \varphi \subseteq M \quad \Rightarrow \quad D_M(\varphi) = D(\varphi)$$

and

$$\operatorname{supp} \varphi \cap M = \emptyset \quad \Rightarrow \quad D_M(\varphi) = 0.$$

(R4) For any family of pairwise disjoint measurable sets $(M_i)_{i\in\mathbb{N}}$ with $M := \bigcup_{i\in\mathbb{N}} M_i$ and any $D \in \mathbb{D}$ the restriction (2.2.1) fulfills

$$D_M = \sum_{i\in\mathbb{N}} D_{M_i},$$

in particular,

$$D_{M_1\cup M_2} = D_{M_1} + D_{M_2}.$$

Furthermore, for any disjoint sets $M_1, M_2 \subseteq \mathbb{R}$ the restriction fulfills

$$(D_{M_1})_{M_2} = 0.$$

Remark 2.2.1 (Support of a restriction)
One might wonder why it is assumed in (R1) that $\operatorname{supp} D_M \subseteq \operatorname{cl} M$ and not $\operatorname{supp} D_M \subseteq M$ for some $M \subseteq \mathbb{R}$. The reason is that the support of a distribution is always a closed set, so one cannot expect that the support of a restriction to an open set will be contained within this open set. However, the second property of (R4) ensures that nothing "essential" will remain on the boundary of M. As an example consider the distribution $\delta_0 + \mathbb{1}_{\mathbb{D}}$, i.e. the sum of the Dirac impulse at zero with one, then a restriction to the open interval $(0,1)$ is $(\mathbb{1}_{(0,1)})_{\mathbb{D}}$ whose support is $[0,1]$, however, (R4) with $M_1 = (0,1)$ and $M_2 = \{0\}$ ensures that the restriction does not contain a distribution with point support at zero. ◻

Theorem 2.2.2 (Distributional restriction impossible)
A distributional restriction (2.2.1) cannot fulfill (R1)-(R4) simultaneously. ◻

For the proof a specific distribution will be used as a counter example, the existence of this distribution will be established first in the following lemma.

Lemma 2.2.3 ("Bad" distribution)
Let $(d_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ be given by $d_n := \frac{(-1)^n}{n+1}$ and let

$$D : \mathcal{C}_0^\infty \to \mathbb{R}, \quad \varphi \mapsto \sum_{i=0}^{\infty} d_n \varphi(d_n).$$

Then $D \in \mathbb{D}$. □

Proof. By the mean value theorem there exists some sequence $(\xi_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ such that

$$\forall n \in \mathbb{N}: \quad \varphi(d_n) = \varphi(0) + d_n\varphi'(\xi_n).$$

Then

$$\sum_{n\in\mathbb{N}} d_n\varphi(d_n) = \varphi(0)\sum_{n\in\mathbb{N}} d_n + \sum_{n\in\mathbb{N}} {d_n}^2\varphi'(\xi_n).$$

By Leibniz' alternating series test, $\sum_{n\in\mathbb{N}} d_n$ exists in $\mathbb{R}$. Since φ' is continuous with compact support, it follows that

$$\left|\sum_{n\in\mathbb{N}} {d_n}^2\varphi'(\xi_n)\right| \leq \sup_{\xi\in\mathbb{R}} |\varphi'(\xi)| \sum_{n\in\mathbb{N}} {d_n}^2 < \infty,$$

which shows that $D(\varphi)$ is well defined.
Invoking Proposition 2.1.8 yields that D is a distribution. $\boxed{\text{qed}}$

Proof (of Theorem 2.2.2). Consider $D \in \mathbb{D}$ from Lemma 2.2.3. It will be shown, that a restriction to $(0,\infty)$ is not possible.
First observe that D can be rewritten as

$$D = \sum_{n\in\mathbb{N}} d_n\delta_{d_n}.$$

Condition (R4) enforces that $D_{(0,\infty)}$ must be the sum of all Dirac impulses with support in $(0,\infty)$, i.e.

$$D_{(0,\infty)} = \sum_{k\in\mathbb{N}} d_{2k}\delta_{d_{2k}}.$$

But now $D_{(0,\infty)}$ is not a distribution any more, because if one considers a test function φ with the property $\varphi(t) = 1$ for all $t \in [0,1]$, then

$$D_{(0,\infty)}(\varphi) = \sum_{k\in\mathbb{N}} \frac{1}{2k+1} = \infty.$$

$\boxed{\text{qed}}$

Instead of trying to define restrictions for all distributions one can try to define an appropriate subspace of $\mathbb{D}$ for which a definition of a restriction with the desired properties is possible. For example, an obvious subspace for which the distributional restriction is well defined is the space of regular distribution (just take (R2) as the definition), but, of course, this would not be satisfactory, because the restriction would not be defined for any "real" distribution.

Definition 2.2.4 (Piecewise-regular distributions)
A set $T \subseteq \mathbb{R}$ is called *locally finite* if, and only if, for all compact sets $K \subseteq \mathbb{R}$ the set $K \cap T$ is finite. The space of *piecewise-regular distributions* is defined as

$$\mathbb{D}_{\text{pwreg}} := \left\{ f_{\mathbb{D}} + \sum_{t \in T} D_t \;\middle|\; \begin{array}{l} f \in L_{1,\text{loc}}, T \subset \mathbb{R} \text{ locally finite,} \\ \forall\, t \in T : D_t \in \mathbb{D}_{\{t\}} \end{array} \right\}. \qquad \square$$

Proposition 2.2.5 (Proper subset)

$$\mathbb{D}_{\text{pwreg}} \subset \mathbb{D}. \qquad \square$$

Proof. From Proposition 2.1.8 it follows that $\mathbb{D}_{\text{pwreg}} \subseteq \mathbb{D}$. To show that $\mathbb{D}_{\text{pwreg}} \neq \mathbb{D}$ it suffices to define a distribution which is not in $\mathbb{D}_{\text{pwreg}}$. This was already done in Lemma 2.2.3. $\boxed{\text{qed}}$

Proposition 2.2.6 (Unique representation)
Let $D \in \mathbb{D}_{\text{pwreg}}$ and assume there exist two locally finite sets $S, T \subseteq \mathbb{R}$, two sets of distributions with point support $\left\{ D_s^S \in \mathbb{D}_{\{s\}} \;\middle|\; s \in S \right\}$ and $\left\{ D_t^T \in \mathbb{D}_{\{t\}} \;\middle|\; t \in T \right\}$, and two locally integrable functions $f, g \in L_{1,\text{loc}}$ with

$$f_{\mathbb{D}} + \sum_{t \in T} D_t^T = D = g_{\mathbb{D}} + \sum_{s \in S} D_s^S.$$

Then

(i) $f = g$ almost everywhere,

(ii) $\forall \tau \in S \cap T : D^S_\tau = D^T_\tau$,

(iii) $\forall s \in S \backslash T : D^S_s = 0$, and $\forall t \in T \backslash S : D^T_t = 0$.

In other words, a piecewise-regular distribution has a "unique" representation. □

Proof. From
$$f_{\mathbb{D}} - g_{\mathbb{D}} = \sum_{s\in S} D^S_s - \sum_{t\in T} D^T_t$$
it follows that the support of the regular distribution $f_{\mathbb{D}} - g_{\mathbb{D}}$ is contained in the countable set $S \cup T$. Hence Proposition 2.1.15 yields $f_{\mathbb{D}} - g_{\mathbb{D}} = 0$, which, together with Proposition 2.1.5, shows $f = g$ almost everywhere.

Let $\tau \in S \cap T$. Seeking a contradiction assume $D^S_\tau(\varphi) \neq D^T_\tau(\varphi)$ for some $\varphi \in \mathcal{C}^\infty_0$. Choose $\varepsilon > 0$ so small that $(\tau - 3\varepsilon, \tau + 3\varepsilon) \cap (S \cup T) = \{\tau\}$ and let $\varphi_\varepsilon := \varphi \mathbb{1}^\varepsilon_{[\tau-\varepsilon,\tau+\varepsilon]}$, where $\mathbb{1}^\varepsilon_{[\tau-\varepsilon,\tau+\varepsilon]}$ is chosen as is in (2.1.1). Then $\operatorname{supp}(\varphi - \varphi_\varepsilon) \subseteq \mathbb{R}\backslash(\tau - \varepsilon, \tau + \varepsilon)$, hence $D^S_\tau(\varphi - \varphi_\varepsilon) = 0 = D^T_\tau(\varphi - \varphi_\varepsilon)$ and therefore $D^S_\tau(\varphi_\varepsilon) \neq D^T_\tau(\varphi_\varepsilon)$. Observe that
$$\sum_{s\in S} D^S_s - \sum_{t\in T} D^T_t = f_{\mathbb{D}} - g_{\mathbb{D}} = 0,$$
which gives the contradiction
$$\sum_{s\in S} D^S_s(\varphi_\varepsilon) = D^S_\tau(\varphi_\varepsilon) \neq D^T_\tau(\varphi_\varepsilon) = \sum_{s\in T} D^T_s(\varphi_\varepsilon).$$
Finally, let $s \in S\backslash T$ and assume $D^S_s(\varphi) \neq 0$ for some $\varphi \in \mathcal{C}^\infty_0$. Defining φ_ε as above such that $\operatorname{supp}\varphi_\varepsilon \cap (S \cup T) = \{s\}$ and $D^S_s(\varphi) = D^S_s(\varphi_\varepsilon)$, this yields the contradiction
$$0 \neq D^S_s(\varphi_\varepsilon) = \sum_{s\in S} D^S_s(\varphi_\varepsilon) = \sum_{t\in T} D^T_t(\varphi_\varepsilon) = 0.$$
For $t \in T\backslash S$ the argument is analogous and omitted. $\boxed{\text{qed}}$

Definition 2.2.7 (Restriction of piecewise-regular distributions)
For $D = f_{\mathbb{D}} + \sum_{t\in T} D_t \in \mathbb{D}_{\text{pwreg}}$ and measurable $M \subseteq \mathbb{R}$, the *restriction of D on M* is defined by

$$D_M := (f_M)_{\mathbb{D}} + \sum_{t\in T} \mathbb{1}_M(t) D_t.$$

□

Proposition 2.2.8 (Restriction well defined)
The restriction

$$\{\ M \subseteq \mathbb{R} \mid M \text{ measurable } \} \times \mathbb{D}_{\text{pwreg}} \to \mathbb{D}_{\text{pwreg}}, \quad (M, D) \mapsto D_M$$

as in Definition 2.2.7 is well defined and fulfills properties (R1)-(R4) for piecewise-regular distributions, i.e. replace $\mathbb{D}$ in (R1)-(R4) by $\mathbb{D}_{\text{pwreg}}$.□

Proof. Proposition 2.2.6 shows that D_M does not depend on the specific representation, hence it remains to show that (R1)-(R4) hold.

(R1) Since for every locally integrable $f \in L_{1,\text{loc}}$ the restriction f_M to some measurable set $M \subseteq \mathbb{R}$ is again a locally integrable function, $(f_M)_{\mathbb{D}}$ is well defined and is a (regular) distribution with support $\operatorname{supp}(f_M)_{\mathbb{D}} \subseteq \operatorname{cl} M$.

To show that $\sum_{t\in T} \mathbb{1}_M(t) D_t$ is a distribution, observe first that for every φ the sum $\left(\sum_{t\in T} \mathbb{1}_M(t) D_t\right)(\varphi)$ is actually a finite sum and is therefore well defined. Now Proposition 2.1.8 ensures that D_M is a distribution.

Clearly, $\operatorname{supp} \sum_{t\in T} \mathbb{1}_M(t) D_t \subseteq M$. Altogether this shows $D_M \in \mathbb{D}_{\operatorname{cl} M}$. Linearity and idempotence follows directly from the definition.

(R2) This property is fulfilled by definition.

(R3) The first implication follows easily from the definition. If $\operatorname{supp}\varphi \cap M = \emptyset$, then $\varphi^{(i)}(t) = 0$ for all $t \in \operatorname{cl} M$ and all $i \in \mathbb{N}$. Hence Proposition 2.1.16 yields $D_M(\varphi) = 0$.

(R4) This property follows directly from the definition and from the fact that $f_M = \sum_{i\in\mathbb{N}} f_{M_i}$ for any family of pairwise disjoint sets $(M_i)_{i\in\mathbb{N}}$ and $M = \bigcup_{i\in\mathbb{N}} M_i$ and for all functions $f : \mathbb{R} \to \mathbb{R}$ (in particular for $f = \mathbb{1}$). $\boxed{\text{qed}}$

Definition 2.2.9 (Impulsive and regular part)
For $t \in \mathbb{R}$ and $D \in \mathbb{D}_{\text{pwreg}}$ with representation $D = f_{\mathbb{D}} + \sum_{t\in T} D_t$ let

$$D[t] := D_{\{t\}} = \begin{cases} D_t, & t \in T, \\ 0, & t \notin T. \end{cases}$$

be the *impulsive part of D at t*. The *impulsive part* of $D \in \mathbb{D}_{\text{pwreg}}$ is defined by

$$D[\cdot] := \sum_{t\in\mathbb{R}} D[t] = \sum_{t\in T} D_t$$

The *regular part* of $D \in \mathbb{D}_{\text{pwreg}}$ is any function $D^{\text{reg}} \in L_{1,\text{loc}}$ such that

$$D_{\text{reg}} := {D^{\text{reg}}}_{\mathbb{D}} = D - D[\cdot] = f_{\mathbb{D}}.$$

A piecewise-regular distribution $D \in \mathbb{D}_{\text{pwreg}}$ is called *impulse free* if, and only if, $D[\cdot] = 0$ or, equivalently, $D = D_{\text{reg}}$. □

The following proposition shows the relationship between the restriction to an *open* interval and the "restriction" to test functions with support in this open set.

Proposition 2.2.10 (Restriction to open interval)
Let $(a, b) \subseteq \mathbb{R}$ be an open interval and let $F, G \in \mathbb{D}_{\text{pwreg}}$. Then the following equivalence holds:

$$F_{(a,b)} = G_{(a,b)} \quad \Leftrightarrow \quad \forall \varphi \in \mathcal{C}_0^\infty \text{ with } \operatorname{supp}\varphi \subseteq (a, b) : F(\varphi) = G(\varphi).$$

□

Proof. Necessity follows from property (R3) of the restriction. To prove sufficiency, let $H := F - G$, then it remains to show that $H_{(a,b)} = 0$ or equivalently $H_{(a,b)}(\varphi) = 0$ for all $\varphi \in \mathcal{C}_0^\infty$. Note that, by assumption,

$$\forall \varphi \in \mathcal{C}_0^\infty : \quad \operatorname{supp}\varphi \subseteq (a, b) \Rightarrow H(\varphi) = H_{(a,b)}(\varphi) = 0.$$

Consider an arbitrary $\varphi \in \mathcal{C}_0^\infty$. For $\varepsilon > 0$ choose $\varphi_l^\varepsilon, \varphi_a^\varepsilon, \varphi_m^\varepsilon, \varphi_b^\varepsilon, \varphi_r^\varepsilon \in \mathcal{C}_0^\infty$ such that

- $\varphi = \varphi_l^\varepsilon + \varphi_a^\varepsilon + \varphi_m^\varepsilon + \varphi_b^\varepsilon + \varphi_r^\varepsilon$,
- $\operatorname{supp} \varphi_l^\varepsilon \subseteq (-\infty, a)$, hence $H_{(a,b)}(\varphi_l^\varepsilon) = 0$,
- $\operatorname{supp} \varphi_a^\varepsilon \subseteq (-\varepsilon + a, a + \varepsilon)$ and $\|\varphi_a^\varepsilon\|_\infty \leq \|\varphi\|_\infty$,
- $\operatorname{supp} \varphi_m^\varepsilon \subseteq (a, b)$, hence $H_{(a,b)}(\varphi_m^\varepsilon) = 0$,
- $\operatorname{supp} \varphi_b^\varepsilon \subseteq (-\varepsilon + b, b + \varepsilon)$ and $\|\varphi_b^\varepsilon\|_\infty \leq \|\varphi\|_\infty$,
- $\operatorname{supp} \varphi_r^\varepsilon \subseteq (b, \infty)$, hence $H_{(a,b)}(\varphi_r^\varepsilon) = 0$.

Then $H_{(a,b)}(\varphi) = H_{(a,b)}(\varphi_a^\varepsilon) + H_{(a,b)}(\varphi_b^\varepsilon)$. Let $\varepsilon > 0$ so small that there are no impulsive parts of $H_{(a,b)}$ in $(a, a+\varepsilon)$ and $(-\varepsilon+b, b)$. Then $H_{(a,b)}(\varphi) = (H_{\mathrm{reg}})_{(a,b)}(\varphi_a^\varepsilon) + (H_{\mathrm{reg}})_{(a,b)}(\varphi_b^\varepsilon)$ and

$$\begin{aligned}
\left|H_{(a,b)}(\varphi)\right| &\leq \int_a^{a+\varepsilon} |H^{\mathrm{reg}}|\,|\varphi_a^\varepsilon| + \int_{-\varepsilon+b}^{b} |H^{\mathrm{reg}}|\,|\varphi_b^\varepsilon| \\
&\leq \|\varphi\|_\infty \left(\int_a^{a+\varepsilon} |H^{\mathrm{reg}}| + \int_{-\varepsilon+b}^{b} |H^{\mathrm{reg}}| \right).
\end{aligned}$$

Since H^{reg} is locally integrable the right hand tends to zero for $\varepsilon \to 0$. This shows that $H_{(a,b)}(\varphi) = 0$. $\boxed{\text{qed}}$

Note that the assertion of Proposition 2.2.10 is in general not true for non-open intervals as is shown in the following example.

Example 2.2.11 (Restriction to non-open intervals)
Consider the interval $[0,1)$ together with $F = \delta_0$ and $G = 0$, then $F(\varphi) = G(\varphi)$ for all $\varphi \in \mathcal{C}_0^\infty$ with $\operatorname{supp}\varphi \subseteq [0,1)$ (because this implies $\varphi(0) = 0$), but $F_{[0,1)} = \delta_0 \neq 0 = G_{[0,1)}$. $\square$

2.2.2 Multiplication with piecewise-smooth functions

Definition 2.2.12 (Piecewise-smooth functions)
The space of *piecewise-smooth functions* is defined by

$$\mathcal{C}^\infty_{\mathrm{pw}} := \left\{ \alpha = \sum_{i \in \mathbb{Z}} \mathbb{1}_{[t_i, t_{i+1})} \alpha_i \;\middle|\; \begin{array}{l} (\alpha_i)_{i\in\mathbb{Z}} \in (\mathcal{C}^\infty)^{\mathbb{Z}}, \\ \{\, t_i \in \mathbb{R} \mid i \in \mathbb{Z} \,\} \text{ locally finite} \\ \text{with } t_i < t_{i+1},\ i \in \mathbb{Z} \end{array} \right\}.$$

□

Note that $\mathcal{C}^\infty_{\mathrm{pw}} \subseteq L_{1,\mathrm{loc}}$.

Remark 2.2.13 (Special properties of $\mathcal{C}^\infty_{\mathrm{pw}}$)
The space of piecewise-smooth functions has two special properties:

(i) Each "smooth piece" of the a piecewise-smooth function is part of a globally smooth function, in particular, there exist functions which are smooth on each interval $[t_i, t_{i+1})$ for a locally finite set $\{\, t_i \mid i \in \mathbb{Z} \,\}$, but which are not piecewise-smooth in the sense of Definition 2.2.12, consider for example the function $\big(t \mapsto \sqrt{|t|}\big)_{(-\infty,0)}$.

(ii) Definition 2.2.12 implies that each piecewise-smooth function is continuous from the right, which seems to be of no significance at this point. However, this leads to the following definition of the multiplication of piecewise-smooth functions and piecewise-regular distribution, where the restrictions of a piecewise-smooth function are "transferred" to restrictions of a piecewise-regular distribution, but for the latter there is an essential difference whether the restriction is $D_{[t_i,t_{i+1})}$ or $D_{(t_i,t_{i+1}]}$. This issue will be discussed further in Section 2.4, see also Remark 2.2.16. □

Definition 2.2.14 (Multiplication with piecewise-smooth functions)
The *multiplication* of a piecewise-regular distribution $D \in \mathbb{D}_{\mathrm{pwreg}}$ with

a piecewise-smooth function $\alpha = \sum_{i\in\mathbb{Z}} \mathbb{1}_{[t_i,t_{i+1})}\alpha_i \in \mathcal{C}^\infty_{\mathrm{pw}}$ is defined by

$$\alpha D = \left(\sum_{i\in\mathbb{Z}} \alpha_{i[t_i,t_{i+1})}\right) D := \sum_{i\in\mathbb{Z}} \alpha_i D_{[t_i,t_{i+1})}.$$

□

Proposition 2.2.15 (Properties of multiplication)
The multiplication as in 2.2.14 is well defined and has the following properties:

(i) $\forall \alpha \in \mathcal{C}^\infty_{\mathrm{pw}}\ \forall D \in \mathbb{D}_{\mathrm{pwreg}}:\ \alpha D \in \mathbb{D}_{\mathrm{pwreg}}$.

(ii) $\forall \alpha,\beta \in \mathcal{C}^\infty_{\mathrm{pw}}\ \forall F,G \in \mathbb{D}_{\mathrm{pwreg}}:\ \alpha(\beta F) = (\alpha\beta)F$, $(\alpha+\beta)F = \alpha F + \beta F$, and $\alpha(F+G) = \alpha F + \alpha G$. □

Proof. Note that the representation $\alpha = \sum_{i\in\mathbb{N}} \mathbb{1}_{[t_i,t_{i+1})}\alpha_i$ is not unique: α_i may vary on $\mathbb{R}\backslash[t_i,t_{i+1})$ without changing α and one can add more points to the set $T = \{\ t_i \mid i\in\mathbb{Z}\ \}$ without changing α. Nevertheless, Corollary 2.1.17 ensures that the term $\alpha_i D_{[t_i,t_{i+1})}$ does not depend on the values of α_i outside the interval $[t_i,t_{i+1})$: if one has another $\widetilde{\alpha_i} \in \mathcal{C}^\infty$ with $\alpha_i(t) = \widetilde{\alpha_i}(t)$ for all $t \in [t_i,t_{i+1})$, then $(\alpha_i - \widetilde{\alpha_i})D_{[t_i,t_{i+1})} = 0$. If T is not "minimal" (i.e. there are points in T at which α is smooth), then property (R4) of the restriction ensures that αD is not changed by the additional points in T. Hence αD does not depend on the specific representation $\alpha = \sum_{i\in\mathbb{N}} \mathbb{1}_{[t_i,t_{i+1})}\alpha_i$ and is therefore well defined.
To show that $\alpha D \in \mathbb{D}_{\mathrm{pwreg}}$, observe first that $\alpha D(\varphi)$ reduces to a finite sum for all test functions $\varphi \in \mathcal{C}^\infty_0$, hence αD is a distribution by Proposition 2.2.8 and Proposition 2.1.8. Finally,

$$\begin{aligned}
\alpha D &= \sum_{i\in\mathbb{Z}} \alpha_i D_{[t_i,t_{i+1})}\\
&= \sum_{i\in\mathbb{Z}} \left(\alpha_i (D_{\mathrm{reg}})_{[t_i,t_{i+1})} + \sum_{t\in[t_i,t_{i+1})} \alpha_i D_t\right)\\
&= (\alpha D^{\mathrm{reg}})_{\mathbb{D}} + \sum_{t\in T} \widetilde{D}_t
\end{aligned}$$

where, by Remark 2.1.13, $\widetilde{D}_t = \alpha_i D_t \in \mathbb{D}_{\{t\}}$ for $i \in \mathbb{Z}$ such that $t \in [t_i, t_{i+1})$. This yields $\alpha D \in \mathbb{D}_{\text{pwreg}}$.
The rest of the proposition follows easily from the definition. $\boxed{\text{qed}}$

Remark 2.2.16 (Restriction is more than multiplication)
The restriction of a piecewise-regular distribution $D \in \mathbb{D}_{\text{pwreg}}$ to an interval of the special form $M = [s, t) \subseteq \mathbb{R}$ can be expressed as a multiplication with a piecewise-smooth function, i.e.

$$\forall\ D \in \mathbb{D}_{\text{pwreg}}\ \forall M = [s, t) : \quad D_M = \mathbb{1}_M D.$$

However, restrictions to other subsets $M \subseteq \mathbb{R}$ cannot be expressed directly as multiplication with $\mathbb{1}_M$ because in general $\mathbb{1}_M \notin \mathcal{C}^\infty_{\text{pw}}$. Since

$$\forall\ s \in \mathbb{R} : \quad D_{\{s\}} = D[s],$$

restriction to other intervals can be expressed as

$$\begin{aligned} D_{(s,t)} &= \mathbb{1}_{[s,t)} D - D[s], \\ D_{(s,t]} &= \mathbb{1}_{[s,t)} D - D[s] + D[t], \\ D_{[s,t]} &= \mathbb{1}_{[s,t)} D + D[t], \end{aligned}$$

where $-\infty \leq s < t \leq \infty$ and $D[\pm\infty] := 0$. □

2.3 Piecewise-smooth distributions and its properties

Definition 2.3.1 (Piecewise-smooth distribution)
The space of *piecewise-smooth distributions* is

$$\mathbb{D}_{\text{pw}\mathcal{C}^\infty} := \left\{ \ D \in \mathbb{D}_{\text{pwreg}} \ \middle|\ \exists f \in \mathcal{C}^\infty_{\text{pw}} : \ D_{\text{reg}} = f_{\mathbb{D}} \ \right\},$$

i.e. a piecewise-regular distribution is called piecewise-smooth if, and only if, its regular part is induced by a piecewise-smooth function. □

Obviously, $\mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ is a linear subspace of $\mathbb{D}_{\text{pwreg}}$. In the following it will be assumed that the regular part $D^{\text{reg}} \in L_{1,\text{loc}}$ of a piecewise-smooth distribution $D \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ is a piecewise-smooth function, in particular $D^{\text{reg}}(t)$ is uniquely defined for every $t \in \mathbb{R}$.

Remark 2.3.2 (Restriction of piecewise-smooth distribtutions)
Since, by definition, each piecewise-smooth distribution is piecewise-regular, the distributional restriction as in Definition 2.2.7 is also well defined for piecewise-smooth distributions and for all measurable sets $M \subseteq \mathbb{R}$. However, a restriction of a piecewise-smooth function to an interval which has not the form $[s, t)$ (e.g. an open interval) is not a piecewise-smooth function in the strict sense of Definition 2.2.12. But since the corresponding regular distribution is invariant under changes of the underlying function on a set of measure zero, the restriction of a piecewise-smooth distribution to any interval (and locally finite unions of intervals) will be a piecewise-smooth distribution again. Nevertheless, a restriction to a general measurable set will not always yield a piecewise-smooth distribution, consider for example the set $M = \bigcup_{n\in\mathbb{N}} \big[1/(2n+1), 1/2n\big)$.

Definition 2.3.3 (Pointwise evaluation)
The *left (right) sided evaluation* of $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ at $t \in \mathbb{R}$ is defined by

$$D(t-) := D^{\mathrm{reg}}(t-) := \lim_{\substack{h\to 0\\ h>0}} D^{\mathrm{reg}}(t-h)$$

and

$$D(t+) := D^{\mathrm{reg}}(t+) = D^{\mathrm{reg}}(t).$$

The *jump* of D at $t \in \mathbb{R}$ is defined as

$$\Delta_t\{D\} := D(t+) - D(t-).$$

□

Note that for every $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$

$$D_{\mathrm{reg}} = \big(t \mapsto D(t-)\big)_{\mathbb{D}} = \big(t \mapsto D(t+)\big)_{\mathbb{D}}.$$

It is worth mentioning that for piecewise-regular distributions $D \in \mathbb{D}_{\mathrm{pwreg}}$ a left or right sided evaluation is in general not possible because $L_{1,\mathrm{loc}}$-functions are not necessarily left or right continuous.

Proposition 2.3.4 (Derivative of a piecewise-smooth distribution) Let $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ with $D^{\mathrm{reg}} = \sum_{i\in\mathbb{Z}} \mathbb{1}_{[t_i,t_{i+1})} f_i$ for some locally finite set $T = \{\ t_i \mid i \in \mathbb{Z}\ \}$ and some smooth $f_i \in \mathcal{C}^\infty$, $i \in \mathbb{Z}$. Then

$$D' = \left(\sum_{i\in\mathbb{Z}} \mathbb{1}_{[t_i,t_{i+1})} f_i{}' \right)_{\mathbb{D}} + \sum_{i\in\mathbb{Z}} \Delta_{t_i}\{D\}\delta_{t_i} + D[\cdot]'. \tag{2.3.1}$$

In particular,

$$D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \quad \Rightarrow \quad D' \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}.$$ □

Proof. By Proposition 2.1.10 it is, for every $i \in \mathbb{Z}$,

$$\begin{aligned}
\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\big(\mathbb{1}_{[t_i,t_{i+1})} f_i\big)_{\mathbb{D}} &= \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\left(f_i\big(\mathbb{1}_{[t_i,t_{i+1})}\big)_{\mathbb{D}}\right)\\
&= f_i{}'\big(\mathbb{1}_{[t_i,t_{i+1})}\big)_{\mathbb{D}} + f_i \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\big(\mathbb{1}_{[t_i,\infty)} - \mathbb{1}_{[t_{i+1},\infty)}\big)_{\mathbb{D}}\\
&= (\mathbb{1}_{[t_i,t_{i+1})} f_i{}')_{\mathbb{D}} + f_i\delta_{t_i} - f_i\delta_{t_{i+1}}\\
&= (\mathbb{1}_{[t_i,t_{i+1})} f_i{}')_{\mathbb{D}} + f_i(t_i)\delta_{t_i} - f_i(t_{i+1})\delta_{t_{i+1}}.
\end{aligned}$$

Now (2.3.1) follows from $f_i(t_i) - f_{i-1}(t_i) = \Delta_{t_i}\{D\}$. Finally, Proposition 2.1.12 implies that $D[\cdot]'$ is again a sum of distributions with point support, hence $D' \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$. $\boxed{\text{qed}}$

Corollary 2.3.5 (Restrictions and derivatives) For all $-\infty \le s \le t \le \infty$ and $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$,

$$\begin{aligned}
\big(D_{[s,t)}\big)' &= (D')_{[s,t)} + D(s-)\delta_s - D(t-)\delta_t,\\
\big(D_{(s,t)}\big)' &= (D')_{(s,t)} + D(s+)\delta_s - D(t-)\delta_t,\\
\big(D_{(s,t]}\big)' &= (D')_{(s,t]} + D(s+)\delta_s - D(t+)\delta_t,\\
\big(D_{[s,t]}\big)' &= (D')_{[s,t]} + D(s-)\delta_s - D(t+)\delta_t,
\end{aligned}$$

where $\delta_{\pm\infty} = 0$. □

Proof. Let $D^{\mathrm{reg}} = \sum_{i\in\mathbb{Z}} \mathbb{1}_{[t_i,t_{i+1})} f_i$ for some $f_i \in \mathcal{C}^\infty$, $i \in \mathbb{Z}$, and some locally finite set $\{\ t_i \in \mathbb{R} \mid i \in \mathbb{Z}\ \}$. Assume, without restriction, that

$s,t \in \{ t_i \mid i \in \mathbb{Z} \}$. From (2.3.1) follows

$$\begin{aligned}
(D_{[s,t)})' - (D')_{[s,t)} &= \sum_{i\in\mathbb{Z}} \Delta_{t_i}\{D_{[s,t)}\}\delta_{t_i} - \left(\sum_{i\in\mathbb{Z}} \Delta_{t_i}\{D\}\delta_{t_i}\right)_{[s,t)} \\
&= \Delta_t\{D_{[s,t)}\}\delta_t - \big(\Delta_s\{D\}\delta_s - \Delta_s\{D_{[s,t)}\}\delta_s\big) \\
&= -D(t-)\delta_t + D(s-)\delta_s.
\end{aligned}$$

This shows the first formula. Since $D[\tau]' = D'[\tau] - \Delta_\tau\{D\}\delta_\tau$ for all $\tau \in \mathbb{R}$ the other three formulae follow easily by linearity of the differential operator and property (R4). qed

Proposition 2.3.6 (Unique distributional antiderivative)
For $D \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ and $t_0 \in \mathbb{R}$ there exists a unique distributional antiderivative

$$H = \int_{t_0} D \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}$$

with $H' = D$ and $H(t_0-) = 0$. Furthermore, for any sequence $(D_n) \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^{\mathbb{N}}$,

$$\int_{t_0} D_n \to \int_{t_0} D \quad \Rightarrow \quad D_n \to D.$$

□

Proof. It follows from Proposition 2.1.18 that every distribution $D \in \mathbb{D}$ has a distributional antiderivative and that all antiderivatives only differ by a constant.
It is first shown, that every distributional antiderivative H of a piecewise-smooth distribution $D = f_{\mathbb{D}} + \sum_{t\in T} D_t \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ is a piecewise-smooth distribution. Let $g : \mathbb{R} \to \mathbb{R}$ be a antiderivative of f, then $g \in \mathcal{C}^\infty_{\text{pw}}$. Every $D_t \in \mathbb{D}_{\{t\}}$ for a fixed $t \in T$ can, by Proposition 2.1.12, be written as

$$D_t = \sum_{i=0}^{n_t} a_t^i \delta_t^{(i)},$$

where $n_t \in \mathbb{N}$ and $a_t^0, \ldots, a_t^{n_t} \in \mathbb{R}$. Clearly, one antiderivative of D_t is given by

$$a_t^0 \big(\mathbb{1}_{[t,\infty)}\big)_{\mathbb{D}} + \sum_{i=1}^{n_t} a_t^i \delta_t^{(i-1)}.$$

Now let

$$h = g + \sum_{t \in T} a_t^0 \mathbb{1}_{[t,\infty)} \in \mathcal{C}^\infty_{\mathrm{pw}}$$

and, for $t \in T$,

$$\widetilde{D}_t = \sum_{i=1}^{n_t} a_t^i \delta_t^{(i-1)} \in \mathbb{D}_{\{t\}},$$

then $H_1 = h_{\mathbb{D}} + \sum_{t \in T} \widetilde{D}_t \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ is a distributional antiderivative of D. Since, by Proposition 2.1.18, all other antiderivatives only differ by a constant, all antiderivatives of D are piecewise-smooth distributions. Let

$$H = H_1 - H_1(t_0-)\mathbb{1}_{\mathbb{D}},$$

then H fulfills $H(t_0-) = 0$ and it is the only antiderivative with this property.
The second assertion follows from Proposition 2.1.8 because $\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} \int_{t_0} D_n = D_n$ and $\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} \int_{t_0} D = D$ $\boxed{\text{qed}}$

2.4 Multiplication of piecewise-smooth distributions

The aim of this section is to define a multiplication for piecewise-smooth distributions. It is shown that there exists a whole family of multiplications which generalize the multiplication of functions, are associative and obey the differentiation rule. However, only two are "time invariant" and can be seen as a "causal" and "anticausal" multiplication. Naturally, the multiplication for piecewise-smooth distributions should generalize the already defined multiplication for piecewise-smooth functions and piecewise-regular distributions as in Definition 2.2.14. This reduces the choices to the causal multiplication.

In view of Remark 2.2.13, it now becomes clear that the choice of piecewise-smooth functions being continuous from the right (which made Definition 2.2.14 intuitive) yield that there is exactly one meaningful multiplication for piecewise-smooth distributions.

2.4.1 Uniqueness of multiplications on $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$

The main ideas of this section stem from the work of Fuchssteiner [Fuc68], see also [Fuc84], where so called *almost bounded* distributions are studied. A distribution $D \in \mathbb{D}$ is called *almost bounded* if, and only if, there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ and $n \in \mathbb{N}$ such that $D = \left(\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\right)^n g_\mathbb{D}$ and, furthermore, for each $k \in \mathbb{N}$ and each finite interval $T \subseteq \mathbb{R}$ there exists a finite set $J_T(k,g) \subseteq T$ such that g is k times uniformly continuously differentiable on $T \setminus J_T(k,g)$. The space of almost bounded distributions is neither a subset of the space of piecewise-smooth distributions nor is it a superset. It is not a subset because the sequence $m_k := |J_T(k,g)|$ corresponding to some $g : \mathbb{R} \to \mathbb{R}$ can grow (locally) unbounded as $k \to \infty$ which yields that the function g is not a piecewise-smooth function as in Definition 2.2.12. It is not a superset because the piecewise-smooth distribution $\sum_{i\in\mathbb{N}} \delta_i^{(i)}$ is not a finite derivative of a continuous function. However, the spaces are very similar and the results for multiplications are identical.

Theorem 2.4.1 (Characterization of all multiplications on $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$)
There exists an algebra $\mathcal{M} : \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \times \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \to \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ with

(M1) $\forall f, g \in \mathcal{C}^\infty_{\mathrm{pw}} : \ \mathcal{M}(f_\mathbb{D}, g_\mathbb{D}) = (fg)_\mathbb{D}$, i.e. the multiplication of functions is generalized,

(M2) $\forall F, G, H \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} : \ \mathcal{M}\big(\mathcal{M}(F,G),H\big) = \mathcal{M}\big(F,\mathcal{M}(G,H)\big)$, i.e. the multiplication is associative,

(M3) $\forall F, G \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} : \ \mathcal{M}(F,G)' = \mathcal{M}(F',G) + \mathcal{M}(F,G')$, i.e. the differentiation rule for a multiplication is fulfilled.

Furthermore, for each algebra $\mathcal{M}$ fulfilling (M1)-(M3) there exists a set $M_\mathcal{M} \subseteq \mathbb{R}$ such that

$$\begin{aligned} \forall t \in M_\mathcal{M} &: \mathcal{M}(\mathbb{1}_{[t,\infty)_\mathbb{D}}, \delta_t) = \delta_t, \\ \forall t \in \mathbb{R}\setminus M_\mathcal{M} &: \mathcal{M}(\mathbb{1}_{[t,\infty)_\mathbb{D}}, \delta_t) = 0, \end{aligned} \tag{2.4.1}$$

and, for two algebras $\mathcal{M}_1, \mathcal{M}_2$ which fulfill (M1)-(M3), the equality of the sets $M_{\mathcal{M}_1} = M_{\mathcal{M}_2}$ implies $\mathcal{M}_1 = \mathcal{M}_2$, i.e. each multiplication satisfying (M1)-(M3) is *uniquely* given by a set $M \subseteq \mathbb{R}$ and (2.4.1). □

Proof. Step 1: All algebras with (M1)-(M3) fulfill (2.4.1)
Let $\mathcal{M}$ be an algebra fulfilling (M1)-(M3). First observe that $0 = \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \mathbb{1}_{(-\infty,t)_{\mathbb{D}}})$ for all $t \in \mathbb{R}$ implies

$$0 = \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \mathbb{1}_{(-\infty,t)_{\mathbb{D}}})' = \mathcal{M}(\delta_t, \mathbb{1}_{(-\infty,t)_{\mathbb{D}}}) + \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, -\delta_t),$$

hence

$$\forall t \in \mathbb{R} : \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t) = \mathcal{M}(\delta_t, \mathbb{1}_{(-\infty,t)_{\mathbb{D}}}). \tag{2.4.2}$$

From Proposition 2.1.10 it follows that, for all smooth $\alpha \in \mathcal{C}^\infty$ and $t \in \mathbb{R}$,

$$\begin{aligned}\mathcal{M}(\alpha_{\mathbb{D}}, \delta_t) &= \mathcal{M}(\alpha_{\mathbb{D}}, \mathbb{1}_{[t,\infty)_{\mathbb{D}}})' - \mathcal{M}(\alpha'_{\mathbb{D}}, \mathbb{1}_{[t,\infty)_{\mathbb{D}}}) \\ &= (\alpha \mathbb{1}_{[t,\infty)})'_{\mathbb{D}} - (\alpha' \mathbb{1}_{[t,\infty)})_{\mathbb{D}} = \alpha\delta_t,\end{aligned}$$

and, inductively for $n \in \mathbb{N}$,

$$\begin{aligned}\mathcal{M}(\alpha_{\mathbb{D}}, \delta_t^{(n+1)}) &= \mathcal{M}(\alpha_D, \delta_t^{(n)})' - \mathcal{M}(\alpha'_{\mathbb{D}}, \delta_t^{(n)}) = (\alpha\delta_t^{(n)})' - \alpha'\delta^{(n)} \\ &= \alpha\delta_t^{(n+1)}.\end{aligned}$$

Altogether this yields

$$\begin{gathered}\forall \alpha \in \mathcal{C}^\infty\ \forall D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} : \mathcal{M}(\alpha_{\mathbb{D}}, D) = \alpha D \\ \text{and analogously } \mathcal{M}(D, \alpha_{\mathbb{D}}) = D\alpha := \alpha D.\end{gathered} \tag{2.4.3}$$

For $t \in \mathbb{R}$ let $H_t := \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t) = \mathcal{M}(\delta_t, \mathbb{1}_{(-\infty,t)_{\mathbb{D}}})$ and $\mathrm{id}_t := (s \mapsto s - t) \in \mathcal{C}^\infty$, then $\mathrm{id}_t\delta_t = 0$ and therefore, by (M2),

$$0 = \mathcal{M}(\mathrm{id}_{t\mathbb{D}}, H_t) = \mathrm{id}_t H_t$$

Now [Jan71, Satz 33.3/4] shows that there exists $a_t \in \mathbb{R}$ such that

$$H_t = a_t\delta_t,$$

furthermore, by (M2),

$$H_t = \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, H_t) = \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, a_t\delta_t) = a_t H_t$$

and it follows that $a_t\delta_t = a_t^2\delta_t$, therefore only the two cases $a_t = 1$ or $a_t = 0$ are possible, which shows (2.4.1).

Step 2: Existence and uniqueness is shown.
It will be shown that (M1)-(M3) together with (2.4.1) uniquely defines a multiplication $\mathcal{M}$. First observe that $\mathcal{M}(F, G)$ for some $F, G \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ can be reduced to a locally finite sum of the following products, $s, t \in \mathbb{R}$, $n, k \in \mathbb{N}$, $\alpha, \beta \in \mathcal{C}^\infty_{\mathrm{pw}}$:

$$\mathcal{M}(\alpha_\mathbb{D}, \beta_\mathbb{D}),\ \mathcal{M}(\alpha_\mathbb{D}, \delta_t^{(n)}),\ \mathcal{M}(\delta_s^{(k)}, \beta_\mathbb{D}),\ \mathcal{M}(\delta_s^{(k)}, \delta_t^{(n)}).$$

The first product is uniquely defined by (M1) hence it remains to study the other three products. Since each piecewise-smooth function $\alpha \in \mathcal{C}^\infty_{\mathrm{pw}}$ can be written as $\alpha = \sum_{i\in\mathbb{Z}} \alpha_i \mathbb{1}_{[t_i,\infty)}$ for some locally finite set $\{\ t_i \in \mathbb{R} \mid i \in \mathbb{Z}\ \}$ and some smooth functions $\alpha_i \in \mathcal{C}^\infty$ and since (2.4.3) holds, the second and third products can further be reduced to locally finite sums of, $s, t \in \mathbb{R}$, $k, n \in \mathbb{N}$,

$$\mathcal{M}(\mathbb{1}_{[s,\infty)_\mathbb{D}}, \delta_t^{(n)}), \mathcal{M}(\delta_s^{(k)}, \mathbb{1}_{[t,\infty)_\mathbb{D}}).$$

Let $s < t$ and choose a smooth function $\alpha_s \in \mathcal{C}^\infty$ such that $\operatorname{supp} \alpha_s \subseteq (-\infty, t)$ and $\alpha_s(\tau) = 1$ for all τ in an neighbourhood of s, then $\alpha_s \delta_s^{(n)} = \delta_s^{(n)}$ for all $n \in \mathbb{N}$. Furthermore, $\mathbb{1}_{[t,\infty)}\alpha_s = 0$, therefore, by (M2),

$$\forall s < t\ \forall n \in \mathbb{N}:\ 0 = \mathcal{M}\Big(\mathbb{1}_{[t,\infty)_\mathbb{D}}, \mathcal{M}\big((\alpha_s)_\mathbb{D}, \delta_s^{(n)}\big)\Big) = \mathcal{M}\big(\mathbb{1}_{[t,\infty)_\mathbb{D}}, \delta_s^{(n)})\big)$$

and, analogously,

$$\begin{aligned}
\forall s < t\ \forall n \in \mathbb{N} : 0 &= \mathcal{M}(\delta_s^{(n)}, \mathbb{1}_{[t,\infty)_\mathbb{D}}),\\
\forall s > t\ \forall n \in \mathbb{N} : 0 &= \mathcal{M}(\mathbb{1}_{(-\infty,t)_\mathbb{D}}, \delta_s^{(n)}),\\
\forall s > t\ \forall n \in \mathbb{N} : 0 &= \mathcal{M}(\delta_s^{(n)}, \mathbb{1}_{(-\infty,t)_\mathbb{D}}).
\end{aligned}$$

Since, for all $s, t \in \mathbb{R}$, $n \in \mathbb{N}$,

$$\mathcal{M}(\mathbb{1}_{[s,\infty)_\mathbb{D}}, \delta_t^{(n)}) = \mathcal{M}((\mathbb{1} - \mathbb{1}_{(-\infty,s)})_\mathbb{D}, \delta_t^{(n)}) = \delta_t^{(n)} - \mathcal{M}(\mathbb{1}_{(-\infty,s)_\mathbb{D}}, \delta_t^{(n)})$$

it also follows that

$$\begin{aligned}
&\forall s > t\ \forall n \in \mathbb{N} : \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_s^{(n)}) = \delta_s^{(n)},\\
&\forall s > t\ \forall n \in \mathbb{N} : \mathcal{M}(\delta_s^{(n)}, \mathbb{1}_{[t,\infty)_{\mathbb{D}}}) = \delta_s^{(n)},\\
&\forall s < t\ \forall n \in \mathbb{N} : \mathcal{M}(\mathbb{1}_{(-\infty,t)_{\mathbb{D}}}, \delta_s^{(n)}) = \delta_s^{(n)},\\
&\forall s < t\ \forall n \in \mathbb{N} : \mathcal{M}(\delta_s^{(n)}, \mathbb{1}_{(-\infty,t)_{\mathbb{D}}}) = \delta_s^{(n)}.
\end{aligned}$$

For $s \neq t$ choose $\alpha_s, \alpha_t \in \mathcal{C}^\infty$ such that $s \notin \operatorname{supp} \alpha_t$, $t \notin \operatorname{supp} \alpha_s$ and that α_s,α_t are constantly one in a neighbourhood of s, t, respectively. Then $\alpha_s \delta_s^{(k)} = \delta_s^{(k)}$, $\alpha_s \delta_t^{(k)} = 0$, $\alpha_t \delta_s^{(k)} = 0$ and $\alpha_t \delta_t^{(k)} = \delta_t^{(k)}$ for all $n, k \in \mathbb{N}$, this implies:

$$\begin{aligned}
\forall s \neq t\ \forall n, k \in \mathbb{N} :\ \mathcal{M}(\delta_s^{(k)}, \delta_t^{(n)}) &= \mathcal{M}(\delta_s^{(k)} \alpha_s, \alpha_t \delta_t^{(n)})\\
&= \mathcal{M}(\delta_s^{(k)} \alpha_t, \alpha_s \delta_t^{(n)}) = 0.
\end{aligned}$$

It remains to study the three products, $t \in \mathbb{R}$, $n, k \in \mathbb{N}$,

$$\mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t^{(n)}),\ \mathcal{M}(\delta_t^{(k)}, \mathbb{1}_{[t,\infty)_{\mathbb{D}}}),\ \mathcal{M}(\delta_t^{(k)}, \delta_t^{(n)}).$$

Consider first $n = k = 0$. The first product is uniquely given by (2.4.1): $\mathcal{M}(\delta_t, \mathbb{1}_{[t,\infty)_{\mathbb{D}}}) = a_t \delta_t$ for a corresponding $a_t \in \{0, 1\}$. The second product follows from (2.4.2) and $\mathcal{M}(\delta_t, \mathbb{1}_{[t,\infty)_{\mathbb{D}}}) = \mathcal{M}(\delta_t, (\mathbb{1} - \mathbb{1}_{(-\infty,t)})_{\mathbb{D}}) = \delta_t - \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t^{(n)}) = (1 - a_t)\delta_t$. From this it follows by (M2) that

$$(1 - a_t)\mathcal{M}(\delta_t, \delta_t) = \mathcal{M}\big(\delta_t, \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t)\big) = a_t \mathcal{M}(\delta_t, \delta_t),$$

hence

$$\mathcal{M}(\delta_t, \delta_t) = 0.$$

Inductively it follows that, for $n, k \in \mathbb{N}$:

$$\begin{aligned}
\mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t^{(n+1)}) &= \mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}}, \delta_t^{(n)})' - \mathcal{M}(\delta_t, \delta_t^{(n)}) = a_t \delta_t^{(n+1)}\\
\mathcal{M}(\delta_t^{(k+1)}, \mathbb{1}_{[t,\infty)_{\mathbb{D}}}) &= \mathcal{M}(\delta_t^{(k)}, \mathbb{1}_{[t,\infty)_{\mathbb{D}}})' - \mathcal{M}(\delta_t^{(k)}, \delta_t) = (1 - a_t)\delta_t^{(k)},
\end{aligned}$$

and

$$\begin{aligned}(1-a_t)\mathcal{M}(\delta_t^{(k+1)},\delta^{(n)}) &= \mathcal{M}(\delta_t^{(k+1)},\mathcal{M}(\mathbb{1}_{[t,\infty)_{\mathbb{D}}},\delta_t^{(n)}) \\ &= a_t\mathcal{M}(\delta_t^{(k+1)},\delta_t^{(n)}),\end{aligned}$$

which implies that $\mathcal{M}(\delta_t^{(k+1)},\delta_t^{(n)}) = 0$. Analogously, it follows that $\mathcal{M}(\delta_t^{(k)},\delta_t^{(n+1)}) = 0$ and $\mathcal{M}(\delta_t^{(k+1)},\delta_t^{(n+1)}) = 0$. This concludes the proof. $\boxed{\text{qed}}$

Corollary 2.4.2 (Unique multiplication on $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$)
There exists a unique multiplication on $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ satisfying (M1)-(M3) and

(M4) $\forall t \in \mathbb{R}: \ \mathbb{1}_{[t,\infty)}\delta_t = \delta_t.$ □

Remark 2.4.3 (Causal and anticausal multiplication)
Under the assumption that a multiplication satisfying (M1)-(M3) is "time-invariant", only two possibilities remain: the multiplications either fulfills (M4) or

(M4') $\forall t \in \mathbb{R}: \ \mathbb{1}_{(-\infty,t)}\delta_t = \delta_t.$

The unique multiplication on $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ satisfying (M1)-(M4) might be called *causal Fuchssteiner multiplication* and the one satisfying (M1)-(M3) and (M4') might be called *anticausal Fuchssteiner multiplication*. The reason for using the term "causal" and "anticausal" is motivated by observing the solution behaviour of the following simple distributional ODE:

$$\dot{x} = \pm\delta_0 x. \tag{2.4.4}$$

Assume first that the causal Fuchssteiner multiplication is used in (2.4.4) with a minus sign. It can be shown that all solutions within $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ of (2.4.4) are given by

$$x = c\mathbb{1}_{(-\infty,0)_{\mathbb{D}}}, \quad c \in \mathbb{R}.$$

Note that $x_{[0,\infty)}$ is identical zero, i.e. for initial value pairs (t_0, x_0) with $t_0 > 0$ only for $x_0 = 0$ there exist solutions, all these solutions

are uniquely determined in the "future" (t_0, ∞) but not in the "past" $(-\infty, t_0)$.
Applying the anticausal Fuchssteiner multiplication to (2.4.4) with a plus sign, the situation is exactly the other way around: All solutions are given by

$$x = c\mathbb{1}_{[0,\infty)_{\mathbb{D}}}, \quad c \in \mathbb{R}$$

and the initial value pair $(t_0, 0)$ with $t_0 < 0$ determines the solution uniquely in the past but not in the future. $\square$

Definition 2.4.4 (Fuchssteiner multiplication)
The *(causal) Fuchssteiner multiplication* is the unique multiplication $\mathcal{M} : \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \times \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \to \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ satisfying (M1)-(M4). For $F, G \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ let $FG := \mathcal{M}(F, G)$. $\square$

In the rest of the work only the causal Fuchssteiner multiplication will be used, therefore the "causal" will be omitted.

Remark 2.4.5 (Square of Dirac impulse)
Let $\delta := \delta_0$, then, as shown in the proof of Theorem 2.4.1,

$$\delta^2 = 0.$$

It is interesting to compare the different approaches to define a multiplication for distributions in the literature with respect to the square of the Dirac impulse: In [Wal94] it is claimed that it is impossible to define this square[1]. A similar result is obtained in [Wal70, Thm. 3.9], however, in the proof it is shown that the square of the Dirac impulse, if it exists, must be zero which contradicts the assumptions made in that paper. In [Mik66] the equation $\delta^2 - \frac{1}{\pi^2}\left(\frac{1}{x}\right)^2 = -\frac{1}{\pi^2}\frac{1}{x^2}$ is established, where the left hand side is considered as a "single entity", this is motivated by quantum mechanics where δ^2 appears only in this context. The square of the Dirac impulse is well defined in [Kön55], but only in a generalized space of distributions and it is shown that δ^2 is not a classical distribution. In [FLZ92] a commutative multiplication for a subspace of distributions is defined and there the square of the Dirac-impulse is zero. $\square$

[1] [Wal94, 3.IV]: "Im besonderen ist es nicht möglich, das Quadrat der δ-Funktion δ^2 zu bilden."

2.4.2 Properties of the Fuchssteiner multiplication

Proposition 2.4.6 (Properties of the Fuchssteiner multiplication) The Fuchssteiner multiplication has the following properties, $F, G \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$, $n \in \mathbb{N}$:

(i) $\delta_t F = F(t-)\delta_t$ and $\delta_t^{(n)} F = \sum_{i=0}^n (-1)^i \binom{n}{i} F^{(i)}(t-)\delta_t^{(n-i)}$,

(ii) $F\delta_t = F(t+)\delta_t$ and $F\delta_t^{(n)} = \sum_{i=0}^n (-1)^i \binom{n}{i} F^{(i)}(t+)\delta_t^{(n-i)}$,

(iii) $F[\cdot]G[\cdot] = 0$,

(iv) in general, $FG \neq GF$,

(v) $\operatorname{supp}(FG) \subseteq \operatorname{supp} F \cap \operatorname{supp} G$.

Proof. (i) The equation $\delta_t F = F(t-)\delta_t$ follows easily from (M4) together with (2.4.3) and an analogon of (2.4.2). The expression for $\delta_t^{(n)} F$ follows by an inductive argument and (M3).

(ii) This follows analogous as above.

(iii) This is an immediate consequence from the above results because $F[\cdot](t+) = 0$ for all $t \in \mathbb{R}$.

(iv) Consider for example the product of $\mathbb{1}_{[0,\infty)_\mathbb{D}}$ and δ_0:

$$\mathbb{1}_{[0,\infty)_\mathbb{D}}\delta_0 = \delta_0 \neq 0 = \delta_0 \mathbb{1}_{[0,\infty)_\mathbb{D}}.$$

(v) This follows from

$$\begin{aligned}
\operatorname{supp} FG &= \operatorname{supp}\big(F^{\mathrm{reg}}G_{\mathrm{reg}} + F^{\mathrm{reg}}G[\cdot] + F[\cdot]G_{\mathrm{reg}} + F[\cdot]G[\cdot]\big) \\
&\subseteq (\operatorname{supp} F_{\mathrm{reg}} \cap \operatorname{supp} G_{\mathrm{reg}}) \cup (\operatorname{supp} F_{\mathrm{reg}} \cap G[\cdot]) \\
&\quad \cup (\operatorname{supp} F[\cdot] \cap \operatorname{supp} G_{\mathrm{reg}}) \cup (\operatorname{supp} F[\cdot] \cap \operatorname{supp} G[\cdot]) \\
&= (\operatorname{supp} F_{\mathrm{reg}} \cup \operatorname{sup} F[\cdot]) \cap (\operatorname{supp} G_{\mathrm{reg}} \cup \operatorname{supp} G[\cdot]) \\
&= \operatorname{supp} F \cap \operatorname{supp} G.
\end{aligned}$$

qed

Remark 2.4.7 (Fuchssteiner multiplication is not "continuous")
Note that the Fuchssteiner multiplication is not "continuous", i.e., for some sequences $(F_n)_{n\in\mathbb{N}}$, $(G_n)_{n\in\mathbb{N}}$ within $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ and some $F, G \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$,

$$F_n \to F \ \wedge\ G_n \to G \quad \not\Rightarrow \quad F_n G_n \to FG.$$

As a simple example consider the sequence $(\delta_{-1/n})_{n\in\mathbb{N}}$ which converges to δ_0. However $\mathbb{1}_{[0,\infty)}\delta_{-1/n} = 0 \neq \delta_0 = \mathbb{1}_{[0,\infty)}\delta_0$. The reason is that already the distributional restriction as in Definition 2.2.7 is not "continuous" in the sense that the restriction of the limit of a sequence need not to be same as limit of the sequence of restriction.
It may be an interesting research topic to define a different convergence for distributions or, equivalently, a different space of test functions, so that the space of "restrictable distributions" (i.e. distributions for which a restriction with properties (R1)-(R4) is possible) can also be understood as a dual space of some space of test functions. □

Proposition 2.4.8 (Multiplication and restriction)
Let $F, G \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ and $s, t \in \mathbb{R} \cup \{\pm\infty\}$ with $s \leq t$, then, for any $\varepsilon > 0$,

$$\begin{aligned}
(FG)_{(s,t)} &= F_{(s,t)}G_{(s,t)} \\
(FG)_{[s,t)} &= F_{[s,t)}G_{[s,t)} + F[s]G_{(s-\varepsilon,s)} \\
(FG)_{(s,t]} &= F_{(s,t]}G_{(s,t]} + F_{(t,t+\varepsilon)}G[t] \\
(FG)_{[s,t]} &= F_{[s,t]}G_{[s,t]} + F_{(t,t+\varepsilon)}G[t] + F[s]G_{(s-\varepsilon,s)}
\end{aligned}$$
□

Proof. Let $M \subseteq \mathbb{R}$ be one of the four intervals with boundaries s and t, then by linearity of the restriction

$$(FG)_M = (F_{\mathrm{reg}}G_{\mathrm{reg}})_M + (F_{\mathrm{reg}}G[\cdot])_M + (F[\cdot]G_{\mathrm{reg}})_M,$$

First observe that $(F_{\mathrm{reg}}G_{\mathrm{reg}})_M = (F_{\mathrm{reg}})_M(G_{\mathrm{reg}})_M$. Furthermore,

$$\begin{aligned}
(F_{\mathrm{reg}}G[\cdot])_M =& \big((F_{\mathrm{reg}})_M G[\cdot]_M\big)_M + \big((F_{\mathrm{reg}})_{\mathbb{R}\setminus M} G[\cdot]_M\big)_M \\
&+ \big(F_{\mathrm{reg}}G[\cdot]_{\mathbb{R}\setminus M})\big)_M,
\end{aligned}$$

where the term $\big((F_{\mathrm{reg}})G[\cdot]_{\mathbb{R}\setminus M})\big)_M$ is zero, because $F_{\mathrm{reg}}G[\cdot]_{\mathbb{R}\setminus M}$ is a distribution with zero regular part and whose support is a locally finite set contained in $\mathbb{R}\setminus M$, hence the restriction to M is zero by definition. Since the support of $(F_{\text{reg}})_M G[\cdot]_M$ is a locally finite set and is contained within M the outer restriction does not change it. Finally, the support of $(F_{\text{reg}})_{\mathbb{R}\setminus M} G[\cdot]_M$ is also a locally finite set and is contained in $\operatorname{cl}(\mathbb{R}\setminus M) \cap M \subseteq \{s,t\}$, hence, if $s<t$,

$$(F_{\text{reg}} G[\cdot])_M = (F_{\text{reg}})_M G[\cdot]_M + (F_{\text{reg}})_{\mathbb{R}\setminus M}\big(G[s]+G[t]\big)_M$$

Analogously,

$$(F[\cdot]G_{\text{reg}})_M = F[\cdot]_M (G_{\text{reg}})_M + \big(F[s]+F[t]\big)_M (G_{\text{reg}})_{\mathbb{R}\setminus M}.$$

Now let $M=(s,t)$, then $\big(G[s]+G[t]\big)_M = 0 = \big(F[s]+F[t]\big)_M$, hence the assertion is shown in this case. For $M=[s,t)$ it is $\big(G[s]+G[t]\big)_M = G[s]$ and $\big(F[s]+F[t]\big)_M = F[s]$. As mentioned in Proposition 2.4.6 the term $(F_{\text{reg}})_{\mathbb{R}\setminus M} G[s]$ depends only on the value $\big((F_{\text{reg}})_{\mathbb{R}\setminus M}\big)^{(i)}(s+)$, $i\in\mathbb{N}$, which is zero for all $i\in\mathbb{N}$, hence $(F_{\text{reg}})_{\mathbb{R}\setminus M} G[s] = 0$. Also from this proposition it follows that $F[s](G_{\text{reg}})_{\mathbb{R}\setminus M} = F[s]G_{(s-\varepsilon,s)}$ for any $\varepsilon>0$. This shows the assertion for $M=[s,t)$. Analogous arguments show the validity of the assertions for $M=(s,t]$ and $M=[s,t]$.
If $s=t$, then

$$\begin{aligned}
(FG)_{[s,t]} &= (FG)[s] = (F_{\text{reg}} G[\cdot])[s] + (F[\cdot]G_{\text{reg}})[s] \\
&= F_{\text{reg}} G[s] + F[s] G_{\text{reg}} \\
&= \underbrace{F[s]G[s]}_{=0} + F_{\text{reg}\,(s,s+\varepsilon)} G[s] + F[s] G_{\text{reg}\,(s-\varepsilon,s)} \\
&= F_{[s,t]} G_{[s,t]} + F_{(t,t+\varepsilon)} G[s] + F[s] G_{(s-\varepsilon,s)}
\end{aligned}$$

$\boxed{\text{qed}}$

2.4.3 Matrix calculus for piecewise-smooth distributions

Definition 2.4.9 (Invertibility of $\mathbb{D}_{\text{pw}\mathcal{C}^\infty}$-matrices)
For two matrices $P \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^{n\times m}$, $Q \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^{m\times p}$, $n,m,p\in\mathbb{N}$, with piecewise-smooth distributional entries the matrix product is defined in the standard way, i.e., for $i=1,\ldots,n$ and $j=1,\ldots,p$,

$$(PQ)_{ij} = \sum_{k=1}^{m} P_{ik} Q_{kj},$$

where M_{ij} denotes the (i,j)-entry of some matrix M. A square matrix $M \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$, $n \in \mathbb{N}$, is called *invertible (over* $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$*)* if, and only if, there exists a matrix $M^{-1} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ such that

$$MM^{-1} = M^{-1}M = I,$$

where $I \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ is the (distributional) identity matrix given by

$$I_{ij} = \begin{cases} \mathbb{1}_{\mathbb{D}}, & i = j \\ 0, & i \neq j \end{cases}.$$

Note that no notational distinction between the matrices $I \in \mathbb{R}^{n\times n}$, $I \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$, and $I \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ is made. □

Proposition 2.4.10 (Invertibility of $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$-matrices)
A piecewise-smooth distributional matrix $M \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$, $n \in \mathbb{N}$, is invertible if, and only if, $M^{\mathrm{reg}} \in \left(\mathcal{C}^\infty_{\mathrm{pw}}\right)^{n\times n}$ is invertible over $\mathcal{C}^\infty_{\mathrm{pw}}$, i.e. there exists $P \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$ with $M^{\mathrm{reg}}(t)P(t) = P(t)M^{\mathrm{reg}}(t) = I$ for all $t \in \mathbb{R}$.
If M is invertible, then the inverse is given by

$$M^{-1} = M_{\mathrm{reg}}^{-1} - M_{\mathrm{reg}}^{-1}M[\cdot]M_{\mathrm{reg}}^{-1}, \quad \text{where } M_{\mathrm{reg}}^{-1} := \big((M^{\mathrm{reg}})^{-1}\big)_{\mathbb{D}}. \qquad \square$$

Proof. If M^{reg} is invertible over $\mathcal{C}^\infty_{\mathrm{pw}}$, then

$$\begin{aligned} MM^{-1} &= (M_{\mathrm{reg}} + M[\cdot])\left(M_{\mathrm{reg}}^{-1} - M_{\mathrm{reg}}^{-1}M[\cdot]M_{\mathrm{reg}}^{-1}\right) \\ &= \underbrace{M_{\mathrm{reg}}M_{\mathrm{reg}}^{-1}}_{=I} \underbrace{-M_{\mathrm{reg}}M_{\mathrm{reg}}^{-1}M[\cdot]M_{\mathrm{reg}}^{-1} + M[\cdot]M_{\mathrm{reg}}^{-1}}_{=0} \\ &\quad - \underbrace{M[\cdot]M_{\mathrm{reg}}^{-1}M[\cdot]M_{\mathrm{reg}}^{-1}}_{=0}. \end{aligned}$$

An analogous calculation shows $M^{-1}M = I$. Hence sufficiency is shown.

Now assume that M is invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$, i.e. there exists a matrix $M^{-1} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ such that $MM^{-1} = I$. Let $M^{-1} = (M^{-1})_{\mathrm{reg}} + M^{-1}[\cdot]$, then

$$\begin{aligned} I = MM^{-1} &= (M_{\mathrm{reg}} + M[\cdot])((M^{-1})_{\mathrm{reg}} + M^{-1}[\cdot]) \\ &= M_{\mathrm{reg}}(M^{-1})_{\mathrm{reg}} + \underbrace{M_{\mathrm{reg}}M^{-1}[\cdot] + M[\cdot](M^{-1})_{\mathrm{reg}}}_{=:H}. \end{aligned}$$

Since $H[\cdot] = H$ and $I[\cdot] = 0$, it follows that H must be zero. This implies

$$I = M_{\mathrm{reg}}(M^{-1})_{\mathrm{reg}} = (M^{\mathrm{reg}}(M^{-1})^{\mathrm{reg}})_{\mathbb{D}},$$

hence M^{reg} is invertible with inverse $(M^{-1})^{\mathrm{reg}} \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$. Finally, from $H = 0$ and the invertibility of M^{reg} it follows that

$$M^{-1}[\cdot] = -(M^{\mathrm{reg}})^{-1}M[\cdot](M^{-1})_{\mathrm{reg}} = -M^{-1}_{\mathrm{reg}}M[\cdot]M^{-1}_{\mathrm{reg}},$$

hence M^{-1} is unique. $\boxed{\text{qed}}$

Remark 2.4.11 (Invertibility over $\mathcal{C}^\infty_{\mathrm{pw}}$)
It is important to note that for the invertibility of $M \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$ over $\mathcal{C}^\infty_{\mathrm{pw}}$ it is not sufficient that $\det M(t) \neq 0$ for all $t \in \mathbb{R}$. As an example consider the 1×1 "matrix" $m(t) = t$ for $t < 0$ and $m(t) = 1$ for $t \geq 0$. Its pointwise inverse is given by $m^{-1}(t) = 1/t$ for $t < 0$ and $m^{-1}(t) = 1$ for $t \geq 0$. Clearly, m^{-1} is not an element of $\mathcal{C}^\infty_{\mathrm{pw}}$ as defined in Definition 2.2.12 because $t \mapsto 1/t$ cannot be extended to a smooth function defined on the whole axis $\mathbb{R}$.
However, if $\det M(t) = \det M(t+) \neq 0$ and $\det M(t-) \neq 0$ for all $t \in \mathbb{R}$, then it is easy to see that $M \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$ is invertible over $\mathcal{C}^\infty_{\mathrm{pw}}$. Another sufficient condition for invertibility over $\mathcal{C}^\infty_{\mathrm{pw}}$ is that $\inf_{t\in\mathbb{R}} \det M(t) > 0$. $\square$

3 Regularity of distributional DAEs

3.1 Initial trajectory problems (ITPs) and DAE-regularity

Definition 3.1.1 (Distributional DAE)
A *distributional DAE* is given by

$$E\dot{x} = Ax + f, \tag{3.1.1}$$

where $E, A \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times n}$ and $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ for $n, m \in \mathbb{N}$. A short hand notation for a DAE in form (3.1.1) is

$$(E, A) \in \Sigma^{m\times n} \quad \text{with inhomogeneity } f.$$

□

Definition 3.1.2 (Solutions of distributional DAEs and ITPs)
Consider a distributional DAE $(E, A) \in \Sigma^{m\times n}$ with inhomogeneity f.

(i) A *(global) solution* of (3.1.1) is a piece-smooth distribution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ for which (3.1.1) is fulfilled. For the multiplications in (3.1.1) the Fuchssteiner multiplication as in Definition 2.4.4 is used, see also Definition 2.4.9. A global solution will also be called *consistent solution* in the following.

(ii) A piecewise-smooth distribution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is called a *local solution* of (3.1.1) on the interval $J \subseteq \mathbb{R}$ if, and only if,

$$(E\dot{x})_J = (Ax + f)_J.$$

(iii) A piecewise-smooth distribution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is called an *ITP solution* with initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ and initial time $t_0 \in \mathbb{R}$ if, and only if, x fulfills the *initial trajectory problem (ITP)*

$$\begin{aligned} (E\dot{x})_{[t_0,\infty)} &= (Ax + f)_{[t_0,\infty)} \\ x_{(-\infty,t_0)} &= x^0_{(-\infty,t_0)}, \end{aligned} \tag{3.1.2}$$

i.e. x is a local solution of (3.1.1) on $[t_0, \infty)$ which coincides with the initial trajectory x^0 on $(-\infty, t_0)$.

□

In the following the regularity definition for classical DAEs will be generalized to distributional DAEs. It is well known that classical regularity is equivalent to 1) the existence of solutions for any (sufficiently smooth) inhomogeneity and 2) the uniqueness of solutions. This characterization of regularity already made it possible to generalize regularity to time-varying DAEs with analytical coefficient matrices [CP83]. In view of the desired application to switched DAEs it is reasonable to assume additionally that there exists a distributional solution for *all*, i.e. consistent and *inconsistent*, initial values. Note that this is not an additional assumption for DAEs with constant coefficient, because it can be shown that all DAEs which are regular in the classical sense have unique (distributional) solution for all initial values.

Definition 3.1.3 (DAE-regularity)
A DAE $(E, A) \in \Sigma^{m\times n}$ is called *DAE-regular* if, and only if, for every inhomogeneity f, for every initial time $t_0 \in \mathbb{R}$ and for every initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ the ITP (3.1.2) has a unique solution. □

Note that there are now two "regularities": One is the regularity of a distribution as in Definition 2.1.4 and the other one is the regularity of a matrix pair as defined above. To avoid confusion the second regularity is called "DAE-regularity".

Examples 3.1.4 (Non-regular DAEs)
There are different reasons why a DAE may not be regular:

(i) Solutions are not uniquely determined by an initial trajectory: for example, the DAE $(E, A) \in \Sigma^{1\times 2}$ with some inhomogeneity f given by

$$[1\ 0]\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} = [0\ 0]\begin{pmatrix}x_1\\ x_2\end{pmatrix} + f,$$

actually reads as $\dot{x}_1 = f$. Hence x_2 can be arbitrary and is not uniquely determined by an initial trajectory.

(ii) Not for all inhomogeneities a solution exists: for example, the DAE $(E, A) \in \Sigma^{2\times 1}$ with inhomgeneity $f = [f_1/f_2]$ given by

$$\begin{bmatrix}1\\0\end{bmatrix}\dot{x} = \begin{bmatrix}0\\0\end{bmatrix}x + \begin{pmatrix}f_1\\ f_2\end{pmatrix}$$

only has solutions if $f_2 = 0$.

(iii) There do not exist ITP solutions even in the homogeneous case, i.e. when $f = 0$ in (3.1.2): for example, consider the DAE $(E, A) \in \Sigma^{1\times 1}$ given by

$$\delta_0 \dot{x} = \delta_0 x$$

and a corresponding ITP with initial time $t_0 = 0$. The ITP (3.1.2) yields that $\dot{x}^0(0-)\delta_0 = x^0(0-)\delta_0$ which implies that the DAE "enforces" the condition $\dot{x}^0(0-) = x^0(0-)$ on the initial trajectory, i.e. not for all initial trajectories an ITP solution exists. □

The next proposition states that DAE-regularity is invariant under certain transformations of the matrix pair $(E, A) \in \Sigma^{m\times n}$. One natural transformation is multiplication of the DAE with an invertible matrix $S \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times m}$ from the left, the other is applying a state transformation $x \mapsto Tz$ for some invertible matrix $T \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$. It is obvious that x is a (global) solution of the DAE (3.1.1) if, and only if, z is a (global) solution of the transformed DAE

$$SET\dot{z} = (SAT - SET')z + Sf.$$

However, it is not immediately clear how this transformation fits together with an ITP, because the initial trajectory and the inhomogeneity must be adapted appropriately.

Proposition 3.1.5 (DAE-regularity and similarity transformations) Let $S \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times m}$ and $T \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ both be invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$. Then $(E, A) \in \Sigma^{m\times n}$ is DAE-regular if, and only if, the transformed DAE $(SET, SAT - SET') \in \Sigma^{m\times n}$ is DAE-regular. □

Proof. First note that it suffices to show one direction of the equivalence, because for $\widetilde{E} := SET$, $\widetilde{A} := SAT - SET'$, $\widetilde{S} := S^{-1}$ and $\widetilde{T} := T^{-1}$ it follows that $(\widetilde{S}\widetilde{E}\widetilde{T}, \widetilde{S}\widetilde{A}\widetilde{T} - \widetilde{S}\widetilde{E}\widetilde{T}') = (E, A)$, where the equation $\widetilde{T}' = -T^{-1}T'T^{-1}$ was used, which itself follows from $0 = (TT^{-1})'$ and the product rule.
For $\widetilde{E}, \widetilde{A}$ as above and $t_0 \in \mathbb{R}$, $\widetilde{x}^0 \in \mathbb{D}^n_{\mathrm{pw}\mathcal{C}^\infty}$, $\widetilde{f} \in \mathbb{D}^m_{\mathrm{pw}\mathcal{C}^\infty}$ it will be shown that every ITP

$$\widetilde{E}\dot{\widetilde{x}} = \widetilde{A}\widetilde{x} + \widetilde{f}, \quad \widetilde{x}_{(-\infty,t_0)} = \widetilde{x}^0_{(-\infty,t_0)}$$

has a unique solution.

Step 1: Existence of a solution.
Let x be the solution of the ITP

$$E\dot{x} = Ax + f, \quad x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)},$$

where

$$f = S^{-1}\widetilde{f}_{[t_0,\infty)} - S^{-1}[t_0]\left(\widetilde{A}\widetilde{x}^0 - \widetilde{E}\dot{\widetilde{x}}^0\right)_{(-\infty,t_0)}$$

and $x^0 = T\widetilde{x}^0$. It will be shown that $\widetilde{x} := T^{-1}x$ is the desired solution. First observe that, by Proposition 2.4.8,

$$\begin{aligned}\widetilde{x}_{(-\infty,t_0)} &= (T^{-1}x)_{(-\infty,t_0)} = T^{-1}_{(-\infty,t_0)}x^0_{(-\infty,t_0)} = T^{-1}_{(-\infty,t_0)}(T\widetilde{x}^0)_{(-\infty,t_0)}\\ &= \widetilde{x}^0_{(-\infty,t_0)}.\end{aligned}$$

Hence it remains to show that

$$(\widetilde{E}\dot{\widetilde{x}})_{[t_0,\infty)} = (\widetilde{A}\widetilde{x})_{[t_0,\infty)} + \widetilde{f}_{[t_0,\infty)},$$

which is equivalent to

$$S^{-1}(\widetilde{E}\dot{\widetilde{x}})_{[t_0,\infty)} = S^{-1}(\widetilde{A}\widetilde{x})_{[t_0,\infty)} + S^{-1}\widetilde{f}_{[t_0,\infty)}.$$

Note that from Proposition 2.4.8 and Property (R4) it follows that, for any $M \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times h}$ and $h = 1$ or $h = n$,

$$\begin{aligned}S^{-1}M_{[t_0,\infty)} &= (S^{-1}M_{[t_0,\infty)})_{(-\infty,t_0)} + (S^{-1}M_{[t_0,\infty)})_{[t_0,\infty)}\\ &= 0 + S^{-1}_{[t_0,\infty)}M_{[t_0,\infty)}\\ &= (S^{-1}M)_{[t_0,\infty)} - S^{-1}[t_0]M_{(-\infty,t_0)}.\end{aligned}$$

Hence $\widetilde{x}$ must fulfill

$$\begin{aligned}&(S^{-1}\widetilde{E}\dot{\widetilde{x}})_{[t_0,\infty)} - S^{-1}[t_0](\widetilde{E}\dot{\widetilde{x}})_{(-\infty,t_0)}\\ &= (S^{-1}\widetilde{A}\widetilde{x})_{[t_0,\infty)} - S^{-1}[t_0](\widetilde{A}\dot{\widetilde{x}})_{(-\infty,t_0)} + (S^{-1}\widetilde{f})_{[t_0,\infty)} - S^{-1}[t_0]\widetilde{f}_{(-\infty,t_0)}.\end{aligned}$$

From $\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}(T^{-1}) = -T^{-1}T'T^{-1}$ it follows that

$$S^{-1}\widetilde{E}\dot{\widetilde{x}} = S^{-1}SET\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}(T^{-1}x) = E\dot{x} - ET'T^{-1}x$$

and
$$S^{-1}\widetilde{A}\widetilde{x} = S^{-1}(SAT - SET')T^{-1}x = Ax - ET'T^{-1}x.$$

Since, by assumption, $(E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)} + f_{[t_0,\infty)}$, it remains to show that

$$\begin{aligned} f_{[t_0,\infty)} = S^{-1}[t_0](\widetilde{E}\dot{\widetilde{x}})_{(-\infty,t_0)} - S^{-1}[t_0](\widetilde{A}\widetilde{x})_{(-\infty,t_0)} + (S^{-1}\widetilde{f})_{[t_0,\infty)} \\ - S^{-1}[t_0]\widetilde{f}_{(-\infty,t_0)}. \end{aligned}$$

Together with Corollary 2.3.5 and Proposition 2.4.8 this follows from

$$\begin{aligned} (\widetilde{E}\dot{\widetilde{x}})_{(-\infty,t_0)} &= \widetilde{E}_{(\infty,t_0)}\dot{\widetilde{x}}_{(\infty,t_0)} \\ &= \widetilde{E}_{(\infty,t_0)}\left(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\left(\widetilde{x}_{(-\infty,t_0)}\right) + \widetilde{x}(t_0-)\delta_{t_0}\right) \\ &= \widetilde{E}_{(\infty,t_0)}\left(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\left(\widetilde{x}^0_{(-\infty,t_0)}\right) + \widetilde{x}^0(t_0-)\delta_{t_0}\right) \\ &= \widetilde{E}_{(\infty,t_0)}\dot{\widetilde{x}}^0_{(\infty,t_0)} = (\widetilde{E}\dot{\widetilde{x}}^0)_{(-\infty,t_0)}, \\ (\widetilde{A}\widetilde{x})_{(-\infty,t_0)} &= (\widetilde{A}\widetilde{x}^0)_{(-\infty,t_0)}, \\ (S^{-1}\widetilde{f})_{[t_0,\infty)} - S^{-1}[t_0]\widetilde{f}_{(-\infty,t_0)} &= S^{-1}_{[t_0,\infty)}\widetilde{f}_{[t_0,\infty)} = S^{-1}\widetilde{f}_{[t_0,\infty)}, \end{aligned}$$

and the definition of f.

Step 2: Uniqueness of a solution.
Let $\widetilde{x}_1$ and $\widetilde{x}_2$ be two solutions of the ITP

$$\widetilde{E}\dot{\widetilde{x}} = \widetilde{A}\widetilde{x} + \widetilde{f}, \quad \widetilde{x}_{(-\infty,t_0)} = \widetilde{x}^0_{(-\infty,t_0)}$$

for some $t_0 \in \mathbb{R}$, $\widetilde{x}^0 \in \mathbb{D}^n_{\mathrm{pw}\mathcal{C}^\infty}$, $\widetilde{f} \in \mathbb{D}^m_{\mathrm{pw}\mathcal{C}^\infty}$. Then $\widetilde{z} := \widetilde{x}_1 - \widetilde{x}_2$ is a solution of the ITP

$$\widetilde{E}\dot{\widetilde{z}} = \widetilde{A}\widetilde{z}, \quad \widetilde{z}_{(-\infty,t_0)} = 0.$$

It will be shown that $z = T\widetilde{z}$ is a solution of the ITP

$$E\dot{z} = Az, \quad z_{(-\infty,t_0)} = 0,$$

it then follows from the DAE-regularity of (E, A) that $z = 0$, hence $\widetilde{z} = 0$ and the uniqueness of solutions is shown.

Clearly, $z_{(-\infty,t_0)} = 0$, hence it remains to show that $(E\dot{z})_{[t_0,\infty)} = (Az)_{[t_0,\infty)}$. It is, by Proposition 2.4.8,

$$\begin{aligned} 0 &= (\widetilde{E}\dot{\widetilde{z}})_{[t_0,\infty)} - (\widetilde{A}\widetilde{z})_{[t_0,\infty)} \\ &= S_{[t_0,\infty)}(E\dot{z} - Az)_{[t_0,\infty)} + S[t_0](E\dot{z} - Az)_{(-\infty,t_0)} \\ &= S(E\dot{z} - Az)_{[t_0,\infty)} + 0, \end{aligned}$$

hence $(E\dot{z})_{[t_0,\infty)} = (Az)_{[t_0,\infty)}$. qed

Remark 3.1.6 (ITP solutions and similarity transformations)
From the proof of Proposition 3.1.5 it becomes clear that $\widetilde{x}$ is a solution of the ITP (3.1.2) for $(\widetilde{E}, \widetilde{A}) := (SET, SAT - SET')$ with initial time t_0, initial trajectory $\widetilde{x}^0$ and inhomogeneity $\widetilde{f}$ if, and only if, $x = T\widetilde{x}$ is a solution of the ITP (3.1.2) with initial time t_0, initial trajectory

$$x^0 = T\widetilde{x}^0$$

and inhomogeneity

$$f = S^{-1}\widetilde{f}_{[t_0,\infty)} - S^{-1}[t_0]\left(\widetilde{A}\widetilde{x}^0 - \widetilde{E}\dot{\widetilde{x}}^0\right)_{(-\infty,t_0)}.$$

□

The next theorem shows that for DAEs with square coefficient matrices any ITP solution corresponds uniquely to a consistent solution of a special "switched" DAE, where the initial trajectory is part of the inhomogeneity. Since in Theorem 3.2.1 it will be shown that all regular DAEs must have square matrix coefficient, it becomes clear that there is a strong relationship between regularity and solvability of DAEs with jumps in the coefficient matrices. In particular, the problem of imposing inconsistent initial values on a DAE can be viewed as finding a consistent solution of a special DAE with jumps in the coefficients.

Theorem 3.1.7 (ITPs are "switched" DAEs)
Let $(E, A) \in \Sigma^{n\times n}$. Then (E, A) is DAE-regular if, and only if, for all $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$, $t_0 \in \mathbb{R}$, and all inhomogeneities f the following DAE has a unique (global) solution

$$E_{\mathrm{itp}}\dot{x} = A_{\mathrm{itp}}x + f_{\mathrm{itp}}, \tag{3.1.3}$$

where $E_{\text{itp}} = E_{[t_0,\infty)}$, $A_{\text{itp}} = I_{(-\infty,t_0)} + A_{[t_0,\infty)}$ and $f_{\text{itp}} = -x^0_{(-\infty,t_0)} + f_{[t_0,\infty)}$. □

Proof. First observe that Proposition 2.4.8 yields

$$\forall F, G \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}\ \forall t_0 \in \mathbb{R}: \quad (FG)_{[t_0,\infty)} = F_{[t_0,\infty)}G.$$

Hence the following equivalences hold:

$$\begin{aligned}
x \text{ solves (3.1.3)} &\Leftrightarrow (E_{\text{itp}}\dot{x})_{(-\infty,t_0)} = (A_{\text{itp}}x)_{(-\infty,t_0)} + (f_{\text{itp}})_{(-\infty,t_0)} \\
&\qquad \wedge\ (E_{\text{itp}}\dot{x})_{[t_0,\infty)} = (A_{\text{itp}}x)_{[t_0,\infty)} + (f_{\text{itp}})_{[t_0,\infty)} \\
&\Leftrightarrow 0 = x_{(-\infty,t_0)} - x^0_{(-\infty,t_0)} \\
&\qquad \wedge\ E_{[t_0,\infty)}\dot{x} = A_{[t_0,\infty)}x + f_{[t_0,\infty)} \\
&\Leftrightarrow x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)} \\
&\qquad \wedge\ (E\dot{x})_{[t_0,\infty)} = (Ax + f)_{[t_0,\infty)} \\
&\Leftrightarrow x \text{ solves the ITP (3.1.2)}
\end{aligned}$$

qed

It should be noted that in most literature on classical DAEs the problem of inconsistent initial values was not motivated in a satisfying way. The underlying problem is: An inconsistent initial value problem is either seen as a special switched DAE (as in Theorem 3.1.7) or interpreted in terms of restrictions (as in Definition 3.1.2(iii)). But multiplication with non-continuous coefficient matrices or restriction to certain intervals is not possible for general distributions hence most approaches in the literature are "vague", because often distributional solutions are considered without specifying the underlying distributional space. One exception is the approach by Rabier and Rheinboldt [RR96b, Thm. 4.1] in the context of impulsive smooth distributions. They defined an initial trajectory problem as follows (translated into the terminology of this dissertation): x is the ITP solution of the DAE (3.1.1) with initial trajectory x^0 and initial time t_0 if, and only if,

$$E\dot{x} = Ax + f_{\text{itp}}, \quad x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)},$$

where $f_{\text{itp}} = (E\dot{x}^0 - Ax^0)_{(-\infty,t_0)} + f_{[t_0,\infty)}$. This approach can be seen as a combination of the two viewpoints presented here, however it has the conceptual disadvantage that the equation is formally overdetermined because on the interval $(-\infty, t_0)$ the solution x must fulfill two equations. Furthermore, it seems unnatural to assume that the original DAE (given primarily by the matrix coefficients) is also valid in the past provided an inconsistent initial value is present. Finally, the Definition 3.1.2(iii), in contrast to the approach in [RR96b], works also for homogeneous DAEs (which play an important role for the stability analysis of DAEs as in Section 4.3).

3.2 Necessary and sufficient conditions for DAE-regularity

Theorem 3.2.1 ($m = n$)
Let $(E, A) \in \Sigma^{m\times n}$ be DAE-regular. Then $m = n$. □

Proof. The proof shows that if $m > n$, then there exists an open interval and an inhomogeneity such that a local solution does not exist. If $m < n$ then it will be shown that the trivial solution for the homogeneous DAE with zero initial trajectory is not unique. So in both cases the DAE can not be regular.

Step 1: $m > n \Rightarrow$ non-existence of local solution.
The main idea is to reduce the original DAE with $m > n$ locally to a smaller DAE, which has a local solution if the original DAE has a local solution. This reduction can be repeated arbitrarily often as long as the original DAE has local solutions, on the other hand a reduction of the size can not be repeated arbitrarily often because of the finite size of the original DAE, hence the assumption that the original DAE always (i.e. for all inhomogeneities) has a local solution can not hold.

Step 1a: Reduction to smaller DAE.
Let $(E, A) \in \Sigma^{m\times n}$ with $m > n$. It will be shown that there exists an open intervals $\tilde{J} \subseteq J \subseteq \mathbb{R}$ and $(\widetilde{E}, \widetilde{A}) \in \Sigma^{\widetilde{m}\times\widetilde{n}}$ with $\widetilde{m} < m$, $\widetilde{n} < n$ and $\widetilde{m} > \widetilde{n}$ such that the following implication holds:

$$\begin{aligned}\forall f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m\, \exists x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n : (E\dot{x})_J = (Ax + f)_J \\ \Rightarrow \quad \forall \widetilde{f} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{\widetilde{m}}\, \exists \widetilde{x} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{\widetilde{n}} : (\widetilde{E}\dot{\widetilde{x}})_{\widetilde{J}} = (\widetilde{A}\widetilde{x} + \widetilde{f})_{\widetilde{J}}.\end{aligned}$$

Choose an open interval $J_1 \subseteq \mathbb{R}$ such that E_{J_1} is impulse and jump free, i.e. there exists a matrix $\overline{E} \in (\mathcal{C}^\infty)^{m\times n}$ such that $E_{J_1} = (\overline{E}_\mathbb{D})_{J_1}$. Choose an open interval $J_2 \subseteq J_1$ such that $t \mapsto \operatorname{rk}\overline{E}(t)$ is constant on J_2. Since $\operatorname{rk}\overline{E}(t) \leq n$ for all $t \in \mathbb{R}$ it follows by Dolezal's Theorem [Dol64] that there exists a matrix function $\overline{S} : J_2 \to \mathbb{R}^{m\times m}$ which is smooth and pointwise invertible such that

$$\overline{S}\,\overline{E} = \begin{pmatrix} E_1 \\ 0 \end{pmatrix} \quad \text{on } J_2,$$

where $E_1 \in (\mathcal{C}^\infty)^{n\times n}$. Choose an open interval $J_3 \subseteq J_2$ such that $\inf_{t\in J_3} \det \overline{S}(t) > 0$, then $S := (I_{\mathbb{R}\setminus J_3} + \overline{S}_{J_3})_\mathbb{D} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times m}$ is invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$.

Let $SA = \left[\begin{smallmatrix} A_1 \\ A_2 \end{smallmatrix}\right]$, where A_1 has size $n \times n$ and A_2 has size $m - n \times n$. Choose an open interval $J \subseteq J_3$ such that there exists $\overline{A}_2 \in (\mathcal{C}^\infty)^{m-n\times n}$ with $(\overline{A}_2)_\mathbb{D} = (A_2)_\mathbb{D}$ on J. If the original DAE (E, A) is locally solvable on J with arbitrary inhomogeneity, then the DAE (SE, SA) must also be solvable on J for arbitrary inhomogeneities, in particular, for each $t \in J$ and $\eta \in \mathbb{R}^{m-n}$ there must exist a $\xi \in \mathbb{R}^n$ such that $0 = \overline{A}_2(t)\xi + \eta$, hence $\overline{A}_2(t)$ must have full row rank $m - n \leq n$ for all $t \in J$. In passing by, note that this implies $m \leq 2n$.

Invoking again Dolezal's Theorem, there exists a pointwise invertible and smooth matrix function $\overline{T} : J \to \mathbb{R}^{n\times n}$ such that $\overline{A}_2\overline{T} = [0, I]$ on J. Choose an open interval $\widetilde{J} \subseteq J$ such that $\inf_{t\in\widetilde{J}} \det T(t) > 0$ then $T := (I_{\mathbb{R}\setminus\widetilde{J}} + \overline{T}_{\widetilde{J}})_\mathbb{D} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ is invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$. Consider the coordinate transformation $z = T^{-1}x$, then local solvability of (E, A) on J for all inhomogeneities implies local solvability of $(SET, SAT - SET')$ on $\widetilde{J} \subseteq J$. By construction,

$$SET = \begin{bmatrix} E_{11} & E_{12} \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad SAT - SET' = \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix} \quad \text{on } \widetilde{J},$$

where $E_{11}, A_{11} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times 2n-m}$.

If $2n = m$, then the new DAE $(SET, SAT - SET')$ restricted on $\widetilde{J}$ reads

$$\begin{aligned} E_{12}\dot{z} &= A_{12}z + f_1, \\ 0 &= z + f_2, \end{aligned}$$

where the inhomogeneity is split into $[f_1, f_2]$ with corresponding sizes. But this implies that $f_1 = A_{12}f_2 - E_{12}f_2'$ hence the inhomogeneity (also for the original DAE) cannot be arbitrary, therefore $2n > m > n$.

In this case the new DAE reads

$$\begin{aligned} E_{11}\dot{z}_1 + E_{12}\dot{z}_2 &= A_{11}z_1 + A_{12}z_2 + f_1, \\ 0 &= z_2 + f_2, \end{aligned}$$

with the corresponding splitting $z = [z_1/z_2]$. Now let $\widetilde{E} := E_{11}$, $\widetilde{A} := A_{11}$, $\widetilde{m} := n < m$, $\widetilde{n} := 2n - m < n = \widetilde{m}$ and

$$\widetilde{f} := f_1 - A_{12}f_2 + E_{12}f_2'.$$

Clearly, local solvability for an arbitrary imhomogeneity of the original DAE (E, A) on J implies local solvability of the reduced DAE $(\widetilde{E}), \widetilde{A})$ on $\widetilde{J} \subseteq J$ for an arbitrary inhomogeneity $\widetilde{f}$, so Step 1a is shown.

Step 1b: Reductio ad absurdum.
The argument of Step 1a can be applied on the reduced DAE $(\widetilde{E}, \widetilde{A})$ such that another reduction is possible. Since the reduction process reduces the size of the corresponding matrices, the reduction cannot be repeated arbitrarily often, on the other hand the reduction process can be applied always when the corresponding DAE has local solutions for arbitrary inhomogeneities. Hence the assumption that $m > n$ and that the original DAE is locally solvable for arbitrary inhomogeneities leads to a contradiction. This concludes Step 1.

Step 2: $m < n \Rightarrow$ trivial solution not unique.
Similar as in Step 1 the DAE will be reduced such that the reduced homogeneous DAE has a locally unique trivial solution if the original homogeneous has a locally unique trivial solution. This reduction can be repeated arbitrarily often which leads to a contradiction.

Step 2a: Reduction to smaller DAE.
Let $(E, A) \in \Sigma^{m\times n}$ with $m < n$. It will be shown that there exist open intervals $\widetilde{J} \subseteq J \subseteq [0, \infty)$ and $(\widetilde{E}, \widetilde{A}) \in \Sigma^{\widetilde{m}\times\widetilde{n}}$ with $\widetilde{m} < m$, $\widetilde{n} < n$ and $\widetilde{m} < \widetilde{n}$ such that following implication holds:

$$\begin{aligned} &\forall x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \text{ with } E\dot{x} = Ax \text{ and } x_{(-\infty,0)} = 0 :\ x_J = 0 \\ &\quad\Rightarrow\quad \forall \widetilde{x} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{\widetilde{n}} \text{ with } \widetilde{E}\dot{\widetilde{x}} = \widetilde{A}\widetilde{x} \text{ and } \widetilde{x}_{(-\infty,0)} = 0 :\ \widetilde{x}_{\widetilde{J}} = 0. \end{aligned}$$

Choose an open interval $J_1 \subseteq [0,\infty)$ such that E_{J_1} is impulse and jump free and let $\overline{E} \in (\mathcal{C}^\infty)^{m\times n}$ be such that $\overline{E}_\mathbb{D} = E$ on J_1. Choose an open interval $J_2 \subseteq J_1$ on which the rank of $\overline{E}$ is constant. Invoking Dolezal's Theorem, choose a pointwise invertible and smooth $\overline{T} : J_2 \to \mathbb{R}^{n\times n}$ such that

$$\overline{E}\,\overline{T} = [E_1 0] \quad \text{on } J_2$$

for some $E_1 \in (\mathcal{C}^\infty)^{m\times m}$. Choose an open interval $J_3 \subseteq J_2$ such that $T := (I_{\mathbb{R}\setminus J} + \overline{T}_{J_3})_\mathbb{D} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ is invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$.

Let $AT - ET' = [A_1, A_2]$, where $A_1 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times m}$ and $A_2 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times (n-m)}$. Choose an open interval $J \subseteq J_3$ and a smooth matrix $\overline{A}_2 \in (\mathcal{C}^\infty)^{m\times(n-m)}$ such that $(\overline{A}_2)_\mathbb{D} = A_2$ on J. Fix an arbitrary $t \in J$ and let $\xi \in \ker \overline{A}_2(t) \subseteq \mathbb{R}^{n-m}$, then it follows that $A_2\xi\delta_t = \overline{A}_2(t)\xi\delta_t = 0$. Therefore, if $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is a solution of the homogeneous DAE (E, A) with $x_{(-\infty,0)} = 0$, then $x_1 := x + T[0/\xi]\delta_t$ is also a solution of the homogeneous DAE (E, A). The assumption that all solution of the homogeneous DAE (E, A) with zero initial condition fulfill $x_J = 0$ now yield $\xi = 0$, hence $\ker \overline{A}_2(t) = \{0\}$. Since $t \in J$ was arbitrary, the column rank of $\overline{A}_2$ must be full on J, i.e. $\mathrm{rk}\overline{A_2}(t) = n - m \leq m$ for all $t \in J$ (in particular, $n \leq 2m$).

Invoking Dolezal's Theorem again, there exists an pointwise invertible and smooth $\overline{S} : J \to \mathbb{R}^{m\times m}$ such that $\overline{SA_2} = [0/I]$ on J. Choose an open interval $\widetilde{J}$ such that $S := (I_{\mathbb{R}\setminus\widetilde{J}} + \overline{S}_{\widetilde{J}})_\mathbb{D} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times m}$ is invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$. By construction

$$SET = \begin{bmatrix} E_{11} & 0 \\ E_{21} & 0 \end{bmatrix} \quad \text{and} \quad SAT - SET' = \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix} \quad \text{on } \widetilde{J}$$

for some $E_{11}, A_{11} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{(2m-n)\times m}$ and $E_{21}, A_{21} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{(n-m)\times m}$.

If $2m = n$, then the DAE $(SET, SAT - SET')$ reads, locally on $\widetilde{J}$,

$$E_{21}\dot{z}_1 = A_{21}z_1 + z_2,$$

where $Tx = z = [z_1/z_2]$. Let $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be a solution of the homogeneous DAE (E, A) with zero initial trajectory, then the above equation implies that $x_1 := x + T^{-1}[z_1/E_{21}\dot{z}_1 - A_{21}z_1]$ for arbitrary $z_1 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$ with $\operatorname{supp} z_1 \subseteq \widetilde{J}$ is also a solutions of the homogeneous

DAE (E, A) with zero initial trajectory. Hence under the assumption that $x_J = 0$ for all such solutions the case $2m = n$ is not possible.

Therefore the homogeneous DAE $(SET, SAT - SET')$ reads, locally on $\widetilde{J}$,

$$\begin{aligned} E_{11}\dot{z}_1 &= A_{11}z_1 \\ E_{21}\dot{z}_1 &= A_{21}z_1 + z_2, \end{aligned}$$

with $z = [z_1/z_2]$ of corresponding size. Since z_2 is uniquely given by z_1 on $\widetilde{J}$ it follows that every solution x of the homogeneous DAE (E, A) with zero initial trajectory fulfills $x_{\widetilde{J}} = 0$ only if all solutions z_1 of the homogeneous DAE $(\widetilde{E}, \widetilde{A}) := (E_{11}, A_{11})$ with zero initial trajectory fulfill $z_1 = 0$ on $\widetilde{J}$. Since $\tilde{n} := m$ and $\tilde{m} := 2m - n < m = \tilde{n}$, the claim of Step 2a is shown.

Step 2b: Reductio ad absurdum.
As in Step 1b, the assumption that the trivial solution of the homogeneous DAE (E, A) with zero initial condition is unique implies that the reduction of Step 2a can be repeated arbitrarily often which is impossible and therefore Step 2 is shown. $\boxed{\text{qed}}$

Proposition 3.2.2 (Sufficient conditions for DAE-regularity)
Let $n \in \mathbb{N}$.

(i) If $(E_0, A_0), (E_1, A_1) \in \Sigma^{n\times n}$ are DAE-regular, then $\big(E_{0(-\infty,t_1)} + E_{1[t_1,\infty)}, A_{0(-\infty,t_1)} + A_{1[t_1,\infty)}\big)$ is also DAE-regular for all $t_1 \in \mathbb{R}$.

(ii) If $(E_i, A_i) \in \Sigma^{n\times n}$, $i \in \mathbb{N}$, is a family of DAE-regular systems and $\{\, t_i \in \mathbb{R} \mid i \in \mathbb{Z} \,\}$ is a locally finite set, then

$$\left(\sum_{i\in\mathbb{Z}} E_{i[t_i,t_{i+1})}, \sum_{i\in\mathbb{Z}} A_{i[t_i,t_{i+1})}\right)$$

is also DAE-regular.

(iii) If $(E, A) \in \Sigma^{n\times n}$ is DAE-regular, then $(E + E_t, A + A_t)$ is also DAE-regular for all $E_t, A_t \in (\mathbb{D}_{\{t\}})^{n\times n}$, $t \in \mathbb{R}$.

(iv) If $(E,A) \in \Sigma^{n\times n}$ is DAE-regular, then $(E+\widetilde{E}[\cdot], A+\widetilde{A}[\cdot])$ is also DAE-regular for all $\widetilde{E},\widetilde{A} \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$. □

Proof. (i) If $t_0 \geq t_1$, then the ITP for (E,A) is identical to the ITP for (E_1,A_1), hence only $t_0 < t_1$ needs to be considered. For $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ and $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ let x^1 be the unique solution of the ITP (E_0,A_0), $x^1_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ with inhomogeneity f and let x be the unique solution of the ITP (E_1,A_1), $x_{(-\infty,t_1)} = x^1_{(-\infty,t_1)}$ with inhomogeneity f.

It will be shown that x is also the unique solution of the ITP (E,A), $x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$. First observe that $x_{(-\infty,t_0)}=x^1_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ because $t_0 < t_1$. Secondly, the following equivalences hold (invoking Proposition 2.4.8)

$$\begin{aligned}
&(E\dot{x})_{[t_0,\infty)} = (Ax+f)_{[t_0,\infty)}\\
\Leftrightarrow\ &(E\dot{x})_{[t_0,t_1)} = (Ax+f)_{[t_0,t_1)}\\
&\quad\wedge\ (E\dot{x})_{[t_1,\infty)} = (Ax+f)_{[t_1,\infty)}\\
\Leftrightarrow\ &(E_0\dot{x}^1)_{[t_0,t_1)} = (A_0x^1+f)_{[t_0,t_1)}\\
&\quad\wedge\ (E_1\dot{x})_{[t_1,\infty)} = (A_1x+f)_{[t_1,\infty)}.
\end{aligned}$$

The last expression is true by the definition of x^1 and x, hence x is a solution of the ITP.

It remains to show that x is unique. Assume that $\widetilde{x}$ is also an ITP solution. Since, by definition, $\widetilde{x}_{(-\infty,t_0)} = x^0_{(-\infty,t_0)} = x_{(-\infty,t_0)}$, it remains to show that $\widetilde{x}_{[t_0,t_1)} = x_{[t_0,t_1)}$ and $\widetilde{x}_{[t_1,\infty)} = x_{[t_1,\infty)}$. Let z and $\widetilde{z}$ be the solutions of the ITP (E_0,A_0), $z_{(-\infty,t_1)} = x_{(-\infty,t_1)}$ and $\widetilde{z}_{(-\infty,t_1)} = \widetilde{x}_{(-\infty,t_1)}$, resp., then

$$(E_0\dot{z})_{[t_0,t_1)} = (E_0\dot{x})_{[t_0,t_1)} = (A_0x+f)_{[t_0,t_1)} = (A_0z+f)_{[t_0,t_1)}$$

and

$$(E_0\dot{\widetilde{z}})_{[t_0,t_1)} = (E_0\dot{\widetilde{x}})_{[t_0,t_1)} = (A_0\widetilde{x}+f)_{[t_0,t_1)} = (A_0\widetilde{z}+f)_{[t_0,t_1)}.$$

Hence z and $\widetilde{z}$ are also solutions of the ITP (E_0, A_0), $z_{(-\infty,t_0)} = x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ and $\widetilde{z}_{(-\infty,t_0)} = \widetilde{x}_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$, resp. Since (E_0, A_0) is DAE-regular it follows that $z = \widetilde{z}$ and therefore $x_{[t_0,t_1)} = \widetilde{x}_{[t_0,t_1)}$. Finally, observe that x and $\widetilde{x}$ are solutions of the ITP (E_1, A_1), $x_{(-\infty,t_1)} = x_{(-\infty,t_1)}$ and $\widetilde{x}_{(-\infty,t_1)} = \widetilde{x}_{(-\infty,t_1)}$, resp. Since (E_1, A_1) is DAE-regular and $x_{(-\infty,t_1)} = \widetilde{x}_{(-\infty,t_1)}$, it follows that $x = \widetilde{x}$.

(ii) Consider the ITP (E, A), $x_{(-\infty,\tau_0)} = \xi^0_{(-\infty,\tau_0)}$ for some initial trajectory ξ^0 and $\tau_0 \in \mathbb{R}$. Without restriction of generality it may be assumed that $t_0 \leq \tau_0 < t_1$ (just by changing the indices). Let x^0 be the solution of the ITP (E_0, A_0), $x^0_{(-\infty,\tau_0)} = \xi^0_{(-\infty,\tau_0)}$ and, for $i \in \mathbb{N}$, let x^{i+1} be the solution of the ITP (E_{i+1}, A_{i+1}), $x^{i+1}_{(-\infty,t_{i+1})} = x^i_{(-\infty,t_{i+1})}$. Then $x = \lim_{i\to\infty} x^i$ is a well defined distribution and it follows by inductively repeating the same arguments as in (i) that x is the unique solution of the ITP (E, A), $x_{(-\infty,\tau_0)} = \xi^0_{(-\infty,\tau_0)}$. Hence (E, A) is DAE-regular.

(iii) Consider the ITP $(E + E_t, A + A_t)$, $x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ for some $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$, $t_0 \in \mathbb{R}$ and with an inhomogeneity $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$. Clearly, if $t_0 > t$ this ITP is identical to the ITP (E, A) with the same initial trajectory and inhomogeneity. Hence it remains to consider $t_0 \leq t$. Let $\widehat{x}$ be the solution of the ITP (E, A), $\widehat{x}_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ with inhomogeneity f and let x be the solution of the ITP (E, A), $x_{(-\infty,t)} = \widehat{x}_{(-\infty,t)}$ with inhomogeneity $\widehat{f} := f + A_t\widehat{x} - E_t\dot{\widehat{x}}$. It will be shown that x is the unique solution of the ITP $(E + E_t, A + A_t)$, $x_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ with inhomogeneity f. First observe that $x_{(-\infty,t_0)} = \widehat{x}_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$. Secondly,

$$\begin{aligned}\big((E + E_t)\dot{x}\big)_{[t_0,t)} &= (E\dot{x})_{[t_0,t)} = (E\dot{\widehat{x}})_{[t_0,t)} = (A\widehat{x} + f)_{[t_0,t)} \\ &= \big((A + A_t)x + f\big)_{[t_0,t)}\end{aligned}$$

and, since $x_{(-\infty,t)} = \widehat{x}_{(-\infty,t)}$,

$$\begin{aligned}
\big((E+E_t)\dot{x}\big)_{[t,\infty)} &= (E\dot{x})_{[t,\infty)} + E_t\dot{x} \\
&= (Ax+\widehat{f})_{[t,\infty)} + E_t\dot{\widehat{x}} \\
&= (Ax+f+A_t\widehat{x})_{[t,\infty)} \\
&= \big((A+A_t)x+f\big)_{[t_0,t)}.
\end{aligned}$$

Hence it remains to show uniqueness of the solution x. Therefore, let $\widetilde{x}$ also be a solution of the ITP $(E+E_t, A+A_t)$, $\widetilde{x}_{(-\infty,t_0)} = x^0_{(-\infty,t_0)}$ with inhomogeneity f. With the same arguments as above it follows that $\widetilde{x}_{(-\infty,t)} = x_{(-\infty,t)} = \widehat{x}_{(-\infty,t)}$. Now

$$\begin{aligned}
&\big((E+E_t)\dot{x}\big)_{[t,\infty)} = \big((A+A_t)x+f\big)_{[t,\infty)} \\
\Leftrightarrow\ &(E\dot{x})_{[t,\infty)} = (Ax+\widehat{f})_{[t,\infty)}
\end{aligned}$$

and the same for $\widetilde{x}$, hence x and $\widetilde{x}$ are both solutions of the ITP (E,A), $x_{(-\infty,t)} = \widehat{x}_{(-\infty,t)}$ with inhomogeneity $\widehat{f}$. Because (E,A) is DAE-regular it follows that $x = \widetilde{x}$.

(iv) Let $T = \{\, t_i \in \mathbb{R} \mid i \in \mathbb{Z} \,\}$ be a locally finite set such that $\widetilde{E}[\cdot] = \sum_{i\in\mathbb{Z}} \widetilde{E}[t_i]$ and $\widetilde{A}[\cdot] = \sum_{i\in\mathbb{Z}} \widetilde{A}[t_i]$. Furthermore, let $E_0 = E$, $A_0 = A$ and, for $k \in \mathbb{N}$, $E_{k+1} = E_k + \widetilde{E}[t_k]$, $E_{-k-1} = E_{-k} + \widetilde{E}[t_{-k}]$, $A_{k+1} = A_k + \widetilde{A}[t_k]$, $A_{-k-1} = A_{-k} + \widetilde{A}[t_{-k}]$. Then it follows inductively from (iii) that (E_i, A_i) is DAE-regular for all $i \in \mathbb{Z}$. Finally,

$$(E+\widetilde{E}[\cdot], A+\widetilde{A}[\cdot]) = \left(\sum_{i\in\mathbb{Z}} E_{i[t_i,t_{i+1})}, \sum_{i\in\mathbb{Z}} A_{i[t_i,t_{i+1})}\right)$$

and regularity follows from (ii). qed

Corollary 3.2.3 (DAE-regularity independent of impulses)
$(E,A) \in \Sigma^{n\times n}$ is DAE-regular if, and only if, $(E_{\mathrm{reg}}, A_{\mathrm{reg}})$ is DAE-regular. □

Remark 3.2.4 (Significance of impulses in coefficient matrices)
The Corollary 3.2.3 does *not* state that the impulses in E and A have no influence on the solutions, in fact, the proof of the Theorem 3.2.2 reveals that the impulsive parts of E and A are preserved in an altered inhomogeneity. In general, the presence of Dirac impulses and its derivatives in E and A yield solutions which might depend also on the derivatives of the initial trajectory. ◻

The following theorem gives necessary conditions for DAE-regularity. The first condition arises by taken successively the derivative of the equation $E\dot{x} = Ax + f$ and to check whether all derivatives of the inhomogeneity can be matched "structurally", i.e. when $x, \dot{x}, \ddot{x}, \ldots$ are seen as independent variables. The resulting matrix is very similar to the so called *derivative array* as in [KM06] which had its origin in [Cam87]. In a distributional framework it is also possible to check, whether the impulsive terms of the inhomogeneity can be matched, this leads to the so called *impulse array*. Note that for time-invariant systems both conditions are equivalent and are actually a characterization of regularity (see Remark 3.2.6).

Theorem 3.2.5 (Necessary conditions for regularity)
Consider a regular DAE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$.

(i) Define the *derivative array of order* $p \in \mathbb{N}$ as the block matrix

$$\mathcal{M}^p \in \left((\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}\right)^{(p+1)\times(p+2)}$$

where each blocks is defined as, for $i = 1, \ldots, p{+}1$, $j = 1, \ldots, p{+}2$,

$$(\mathcal{M}^p)_{i,j} = \tbinom{i-1}{j-2} E^{(i-j+1)} - \tbinom{i-1}{j-1} A^{(i-j)},$$

with the convention that $\binom{0}{0} = 1$ and $\binom{n}{-k} = \binom{n}{n+k} = 0$ for $k > 0$, $n \in \mathbb{N}$, i.e.

$$\mathcal{M}^p = \begin{bmatrix} -A & E & & & & \\ -A' & E' - A & E & & & \\ -A'' & E'' - 2A' & 2E' - A & E & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ -A^{(p)} & E^{(p)} - pA^{(p-1)} & pE^{(p-1)} - \binom{p}{2} A^{(p-2)} & \binom{p}{2} E^{(p-2)} - \binom{p}{3} A^{(p-3)} & \cdots & E \end{bmatrix}.$$

Then $\mathcal{M}^p(t+)$ and $\mathcal{M}^p(t-)$ have full row rank for all $p \in \mathbb{N}$ and $t \in \mathbb{N}$.

(ii) Define the *impulse array of order* $p \times P$, $p, P \in \cup\mathbb{N}$, as the block matrix

$$\mathcal{N}^{p,P} \in \left((\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}\right)^{(p+2)\times(P+2)}$$

where, for $i = 1, \ldots, p+1$, $j = 1, \ldots, P+1$,

$$(\mathcal{N}^{p,P})_{i,j} = (-1)^i \left(\tbinom{j-1}{j-i} E^{(j-i)} - \tbinom{j-2}{j-i-1} A^{(j-i-1)}\right),$$

with the convention that $\binom{0}{k} = \binom{-1}{k} = 0$ for all $k \in \mathbb{Z}$, i.e.

$$\mathcal{N}^{p,P} = \begin{bmatrix} E & E'-A & E''-A' & \cdots & E^{(P)} - A^{(P-1)} \\ & -E & -2E'+A & \cdots & -PE^{(P-1)} + (P-1)A^{(P-2)} \\ & & \ddots & & \vdots \\ & & (-1)^p E & \cdots & (-1)^p\left(\tbinom{P}{p}E^{P-p} - \tbinom{P-1}{P-p-1}A^{(P-p-1)}\right) \end{bmatrix}.$$

Then for all $p \in \mathbb{N}$ there exists $P \in \mathbb{N}$ such that $\mathcal{N}^{p,P}(t+)$ has full row rank for all $t \in \mathbb{R}$. □

Proof. Let $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be an arbitrary inhomogeneity.

(i) Taking successively the derivative of the equation $E\dot{x} = Ax + f$ yields

$$\begin{aligned} E\dot{x} - Ax &= f \\ E\ddot{x} + (E' - A)\dot{x} - A'x &= f' \\ E\dddot{x} + (2E' - A)\ddot{x} + (E'' - 2A')\dot{x} - A''x &= f'' \\ &\vdots \end{aligned}$$

and it follows inductively that, for all $p \in \mathbb{N}$,

$$\mathcal{M}^p \begin{pmatrix} x \\ \dot{x} \\ \vdots \\ x^{(p)} \\ x^{(p+1)} \end{pmatrix} = \begin{pmatrix} f \\ f' \\ \vdots \\ f^{(p)} \end{pmatrix}$$

and, in particular, for all $t \in \mathbb{R}$,

$$\mathcal{M}^p(t\pm)\begin{pmatrix} x(t\pm) \\ \dot{x}(t\pm) \\ \vdots \\ x^{(p)}(t\pm) \\ x^{(p+1)}(t\pm) \end{pmatrix} = \begin{pmatrix} f(t\pm) \\ f'(t\pm) \\ \vdots \\ f^{(p)}(t\pm) \end{pmatrix}$$

Since (E, A) is assumed to be DAE-regular there exists a solution for any given inhomogeneity f, hence $\mathcal{M}^p(t+)$ and $\mathcal{M}^p(t-)$ must both have full row rank.

(ii) For a fixed $t_0 \in \mathbb{R}$ consider the impulsive part of the DAE (3.1.1) at t_0:

$$(E\dot{x})[t_0] = (Ax + f)[t_0]$$

or, equivalently,

$$\begin{aligned} E_{(t_0,\infty)}\dot{x}[t_0] - A_{(t_0,\infty)}x[t_0] &= A[t_0]x_{(-\infty,t_0)} - E[t_0]\dot{x}_{(-\infty,t_0)} + f[t_0] \\ &=: \widetilde{f}[t_0]. \end{aligned}$$

Since $\widetilde{f}[t_0]$ can be assumed to be arbitrary, and since (E, A) is DAE-regular it follows that the operator

$$(E\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} - A)_{t_0} : (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \to \mathbb{D}^n_{\{t_0\}}, \quad x \mapsto E_{(t_0,\infty)}\dot{x}[t_0] - A_{(t_0,\infty)}x[t_0]$$

must be surjective. Assume

$$x[t_0] = \sum_{i=0}^{p} x_i \delta_{t_0}^{(i)}$$

for some $p \in \mathbb{N}$ and $x_0, x_1, \ldots, x_p \in \mathbb{R}^n$, then

$$\dot{x}[t_0] = \sum_{i=-1}^{p} x_i \delta_{t_0}^{(i+1)},$$

where $x_{-1} := x(t+) - x(t-)$. Easy calculations (using Proposition 2.4.6) yield

$$E_{(t_0,\infty)}\dot{x}[t_0]-A_{(t_0,\infty)}x[t_0] = a_0\delta_{t_0}+a_1\delta'_{t_0}+\cdots+a_{p+1}\delta_{t_0}^{(p+1)} \overset{!}{=} \widetilde{f}[t_0],$$

where

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{p+1} \end{pmatrix} = \mathcal{N}^{p+1,p+1}(t_0+) \begin{pmatrix} x_{-1} \\ x_0 \\ \vdots \\ x_p \end{pmatrix}.$$

Note that, in particular, for $i = 0, 1, \ldots, p+1$

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_i \end{pmatrix} = \mathcal{N}^{i,p+1}(t_0+) \begin{pmatrix} x_{-1} \\ x_0 \\ \vdots \\ x_p \end{pmatrix}.$$

Since $\widetilde{f}[t_0]$ is arbitrary it follows that $a_0, a_1, \ldots,$ can be arbitrary. Hence surjectivity of the operator $(E\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t} - A)_{t_0}$ implies that for all $i \in \mathbb{N}$ there must exist $P_i \in \mathbb{N}$ such that $\mathcal{N}^{i,P_i}$ has full row rank. This proves the theorem. $\boxed{\text{qed}}$

Remark 3.2.6 (Necessary condition for constant coefficient case) Applying Theorem 3.2.5 to the constant coefficient case both conditions reduces to the simple condition that all matrices

$$\begin{bmatrix} -E & A & & & & \\ & -E & A & & & \\ & & -E & A & & \\ & & & \ddots & \ddots & \\ & & & & -E & A \end{bmatrix}$$

have full row rank. Actually, this condition is a characterization of classical regularity of time-invariant DAEs [YS81]. $\square$

Example 3.2.7 (Regularity in the sense of [RR96a])
Consider the case of a DAE with analytical coefficients, then, as already mentioned in the introduction, DAE-regularity implies in particular that $[-E(t), A(t)]$ must have full rank and hence regularity in the sense of [RR96a] is implied. However, consider the DAE $t\dot{x}(t) = x(t)$ which is (completely) regular in the sense of [RR96a], then all solutions are given by

$$x(t) = \begin{cases} \alpha t, & t < 0, \\ 0, & t = 0, \\ \beta t, & t > 0, \end{cases}$$

where $\alpha, \beta \in \mathbb{R}$. In particular, the absolute continuous solution is not uniquely given by the past, because β can be chosen independently from α, hence the example is not DAE-regular. For more examples of this type see [IM05]. ◻

3.3 Distributional ODEs

3.3.1 Consistent solutions of distributional ODEs

Definition 3.3.1 (Distributional ODE)
A DAE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$, is called a *distributional ODE* if, and only if, E is invertible over $\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$. A distributional ODE is called in *standard form* if, and only if, $E = I$. ◻

Consider a distributional ODE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$, with some inhomogeneity $f \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$, then x solves $E\dot{x} = Ax + f$ if, and only if, x solves $\dot{x} = E^{-1}Ax + E^{-1}f$. Hence for the purpose of finding consistent solutions, it suffices to consider in the following the distributional ODE in standard form

$$\dot{x} = Ax + f. \tag{3.3.1}$$

The aim of this section is to show that there exists a solution formula similar to the classical ODE case. More precisely, it is shown that if there exists $t_0 \in \mathbb{R}$ such that $A[\cdot]_{(-\infty,t_0)} = 0$, then there exists a transition matrix Φ_{t_0} and a linear operator Ψ_{t_0} such that every solution x of the distributional ODE (3.3.1) can be written as

$$x = \Phi_{t_0} x_0 + \Psi_{t_0}(f), \quad x_0 \in \mathbb{R}^n.$$

In other words, the solution x can be decomposed into a free motion and a forced (by f) motion. The condition that the coefficient matrix A must be impulse free in the past is necessary to avoid problems which will be discussed in Remark 3.3.4. See also Section 3.3.3.

Before stating the main result several technical lemmas are needed. The next lemma studies the fundamental solution of a classical ODE, here the focus is on the piecewise-smooth properties (as in Definition 2.2.12) of the fundamental solution and its inverse.

Lemma 3.3.2 (Fundamental solution of classical ODE)
Let $\widehat{A} \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$, $n \in \mathbb{N}$. Then, for every $t_0 \in \mathbb{R}$ there exists a unique matrix $\phi(\cdot, t_0) \in \mathcal{C}^\infty_{\mathrm{pw}}$ which is absolutely continuous and fulfills:

$$\phi(\cdot, t_0)' = \widehat{A}\phi(\cdot, t_0) \text{ almost everywhere and } \phi(t_0, t_0) = I. \tag{3.3.2}$$

Furthermore,

$$\forall s, t, t_0 \in \mathbb{R} : \quad \phi(s, t_0) = \phi(s, t)\phi(t, t_0),$$

and $\phi(\cdot, t_0)$ is invertible over $(\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$:

$$\forall t, t_0 \in \mathbb{R} : \phi(t, t_0)^{-1} = \phi(t_0, t).$$

The matrix $\phi(\cdot, t_0)$ is called *fundamental solution* of the (classical) ODE $x' = \widehat{A}x$. □

Proof. The existence of an absolutely continuous fundamental solution $\phi(\cdot, t_0)$, $t_0 \in \mathbb{R}$, is shown in [Son98, C.4], there it is also shown (Remark C.4.3) that $\phi(\cdot, t_0)$ is smooth on intervals and has the above properties.

It remains to show that $\phi(\cdot, t_0)$ is piecewise-smooth as defined in Definition 2.2.12. Let $T = \{ \tau_i \in \mathbb{R} \mid i \in \mathbb{Z} \}$ be a locally finite set such that $\widehat{A} = \sum_{i\in\mathbb{Z}} A_{i[\tau_i, \tau_{i+1})}$ for some family of smooth matrices $(A_i)_{i\in\mathbb{Z}}$. For $t_0 \in \mathbb{R}$ and $i \in \mathbb{Z}$ let $\phi_i(\cdot, t_0)$ be the fundamental solutions of $\dot{x} = A_i x$. Then $\phi_i(\cdot, t)$ is smooth for all $i \in \mathbb{Z}$ because each A_i is smooth. Since the ODEs $\dot{x} = \widehat{A}x$ and $\dot{x} = A_i x$ are identical on the interval $[\tau_i, \tau_{i+1})$, the fundamental solution restricted to this interval are also identical if the initial time fulfills $t \in [\tau_i, \tau_{i+1})$, hence $\phi(s, t) =$

$\phi_i(s,t)$ for all $s,t \in [\tau_i, \tau_{i+1})$. For a fixed $t_0 \in \mathbb{R}$ this yields $\phi(t,t_0) = \phi_i(t,\tau_i)\phi(\tau_i,t_0)$ where $i \in \mathbb{Z}$ is chosen such that $t_0 \in [\tau_i, \tau_{i+1})$. Now it follows that

$$\phi(\cdot,t_0) = \sum_{i\in\mathbb{Z}} \big(\phi_i(\cdot,\tau_i)\phi(\tau_i,t_0)\big)_{[\tau_i,\tau_{i+1})},$$

which shows that $\phi(\cdot,t_0)$ as well is piecewise-smooth. Since $\phi_i(\cdot,\tau_i)$ is invertible over $\mathcal{C}^\infty$ it also follows that the inverse

$$\phi(\cdot,t_0)^{-1} = \sum_{i\in\mathbb{Z}} \big(\phi(t_0,\tau_i)\phi_i(\cdot,\tau_i)^{-1}\big)_{[\tau_i,\tau_{i+1})},$$

is a piecewise-smooth matrix-function. $\boxed{\text{qed}}$

The following lemma is used later several times to show uniqueness of solutions.

Lemma 3.3.3 (Unique trivial solution)
Let $t_0 \in \mathbb{R}$, $A \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$, $n \in \mathbb{N}$ and assume $A[\cdot]_{(-\infty,t_0)} = 0$. Then only the trivial solution $x = 0$ fulfills

$$\dot{x} = Ax, \quad x(t_0-) = 0, \tag{3.3.3}$$

$\square$

Proof. Consider first some open interval $(a,b) \subseteq \mathbb{R}$ for which $A_{(a,b)}$ is impulse free. Let x be any solution of (3.3.3) and define $\eta := \big(\phi(\cdot,\tau)^{-1}\big)_{\mathbb{D}} x$, where $\phi(\cdot,\tau)$ is the fundamental solution of the classical ODE $\dot{x} = A^{\mathrm{reg}}x$ with initial time $\tau \in \mathbb{R}$ as in Lemma 3.3.2. Then, since $x = \phi(\cdot,\tau)\eta$,

$$\dot{x} = A_{\mathrm{reg}}\phi(\cdot,\tau)\eta + \phi(\cdot,\tau)\dot{\eta},$$

hence

$$\dot{\eta} = \phi(\cdot,\tau)^{-1}A[\cdot]x$$

and, in particular,

$$(\dot{\eta})_{(a,b)} = 0.$$

Using [Wal94, 6.II.Cor.] together with Proposition 2.2.10 yields that

$$\eta_{(a,b)} = C_{(a,b)}$$

where $C = (t \mapsto c)_{\mathbb{D}} \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$, $c \in \mathbb{R}^n$, is any constant distribution. As a consequence this yields for all solutions x of (3.3.3) and for all open intervals $(a,b) \subseteq \mathbb{R}$ for which $A_{(a,b)}$ is impulse free that

$$\forall\, \tau \in \mathbb{R}\ \exists c \in \mathbb{R}^n: \quad x_{(a,b)} = (\phi(\cdot,\tau)_{\mathbb{D}})_{(a,b)} c. \tag{3.3.4}$$

Let $T = \{\ t_i \in \mathbb{R}\ \mid i \in \mathbb{N}\ \}$ with $t_0 < t_1 < t_2 < \ldots$ such that $A[t] = 0$ for all $t \notin T$. By induction it will be shown that $x_{(-\infty,t_i)} = 0$ for all $i \in \mathbb{N}$. From (3.3.4) it follows that

$$x_{(-\infty,t_0)} = (\phi(\cdot,t_0)_{\mathbb{D}})_{(-\infty,t_0)} c$$

for some $c \in \mathbb{R}^n$. Since $x(t_0-) = 0$ and $\phi(\cdot,t_0)_{\mathbb{D}}(t_0-) = \phi(t_0,t_0) = I$ it follows that $c = 0$, hence

$$x_{(-\infty,t_0)} = 0.$$

Assume now that $x_{(-\infty,t_i)} = 0$ for some $i \in \mathbb{N}$. Then

$$\dot{x}[t_i] = (Ax)[t_i] = A[t_i]x + A_{\mathrm{reg}} x[t_i],$$

since $x^{(j)}(t_i-) = 0$ for all $j \in \mathbb{N}$ it follows from the definition of the product that $A[t_i]x = 0$, hence

$$\dot{x}[t_i] = A_{\mathrm{reg}} x[t_i].$$

Seeking a contradiction assume $x[t_i] \neq 0$, i.e. $x[t_i] = \sum_{k=0}^{n_i} a_i \delta_{t_i}^{(k)}$, $n_i \in \mathbb{N}$, for some $a_1, \ldots, a_{n_i} \in \mathbb{R}^n$ and with $a_{n_i} \neq 0$, then

$$\dot{x}[t_i] = \big(x(t_i+) - x(t_i-)\big)\delta_{t_i} + \sum_{k=0}^{n_i} a_i \delta_{t_i}^{(k+1)} \overset{!}{=} A_{\mathrm{reg}} x[t_i].$$

Since $A_{\mathrm{reg}} x[t_i]$ contains no term $\delta_{t_i}^{(n_i+1)}$ the coefficient vector a_{n_i} must be zero (see also Remark 2.1.14), which contradicts the assumption, hence $x[t_i] = 0$ and therefore $\dot{x}[t_i] = A_{\mathrm{reg}}x[t_i] = 0$. In particular, $x(t_i+) = x(t_i-) = 0$. Invoking (3.3.4) with $\tau = t_i$ and $x(t_i+) = 0$ yields $x_{(t_i,t_{i+1})} = 0$. Altogether this shows $x_{(-\infty,t_{i+1})} = x_{(-\infty,t_i)} + x[t_i] + x_{(t_i,t_{i+1})} = 0$ and the proof is complete. $\boxed{\text{qed}}$

Remark 3.3.4 (Impulses in the past)
If the condition $A[\cdot]_{(-\infty,t_0)} = 0$ is not fulfilled, then, in general, the assertion of Lemma 3.3.3 is not true. This can be seen by the simple distributional ODE $\dot{x} = -\delta_0 x$ which has the solutions $x = c\mathbb{1}_{(-\infty,0)}$, $c \in \mathbb{R}$, hence, for $t_0 > 0$ the condition $x(t_0-) = 0$ does not imply $x = 0$, see also Remark 2.4.3. The underlying problem is that it is, in general, not possible to solve distributional ODEs “backward” in time. $\square$

The matrix Φ_{t_0} and the linear operator Ψ_{t_0} in the desired solution formula $x = \Phi_{t_0}x_0 + \Psi_{t_0}(f)$, $x_0 \in \mathbb{R}$ will be given by the limits of certain sequences. In the following lemma the existence of a general family of limits is shown.

Lemma 3.3.5 (Convergence of a special sequence of distributions)
Let $H, G \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$, $n \in \mathbb{N}$, with $H_{\mathrm{reg}} = 0$ and $H_{(-\infty,t_0)} = 0$ for some $t_0 \in \mathbb{R}$. Let $(S_i)_{i\in\mathbb{N}} \in \left((\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times m}\right)^{\mathbb{N}}$, $m \in \mathbb{N}$, be some sequence such that, for all $i \in \mathbb{N}$,

$$S_{i+1} - S_i = G\int_{t_0} H(S_i - S_{i-1}),$$

where $S_{-1} := 0$. Then $(S_i)_{i\in\mathbb{N}}$ converges to some $S \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times m}$ as $i \to \infty$ and for all $t \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that, for all $i \geq N$,

$$S_{(-\infty,t)} = (S_i)_{(-\infty,t)}. \qquad \square$$

Proof. By assumption, $H = \sum_{i=1}^{\infty} H[t_i]$ for some $t_1, t_2, \ldots \in \mathbb{R}$ with $t_0 \leq t_1 < t_2 < \ldots$ and $t_i \to \infty$ for $i \to \infty$.

Step 1: It is shown that $\operatorname{supp}(S_{i+1} - S_i) \subseteq [t_{i+1}, \infty)$ *for all* $i \in \mathbb{N}$. This assertion is shown by induction. By definition

$$S_1 - S_0 = G\int_{t_0} HS_0,$$

and by Proposition 2.4.6(v) it follows that

$$\operatorname{supp} HS_0 \subseteq \operatorname{supp} H = \{t_1, t_2, t_3, \ldots\}.$$

Since $t_0 \leq t_1$,

$$\operatorname{supp} \int_{t_0} HS_0 \subseteq [t_1, \infty)$$

and invoking again Proposition 2.4.6(v) yields

$$\operatorname{supp}(S_1 - S_0) \subseteq [t_1, \infty).$$

Now assume $\operatorname{supp}(S_i - S_{i-1}) \subseteq [t_i, \infty)$ for some $i \in \mathbb{N}$, then $H[t_i](S_i - S_{i-1}) = 0$. Together with Proposition 2.4.6(v) this implies

$$\operatorname{supp} H(S_i - S_{i-1}) \subseteq \{t_{i+1}, t_{i+2}, \ldots\},$$

which analogously as above yields that $\operatorname{supp} \int_{t_0} H(S_i - S_{i-1}) \subseteq [t_{i+1}, \infty)$ and

$$\operatorname{supp}(S_{i+1} - S_i) \subseteq [t_{i+1}, \infty).$$

This concludes Step 1.

Step 2: Convergence of the sequence (S_i) *is shown.*
For every test function $\varphi \in \mathcal{C}_0^\infty$ there exists $N \in \mathbb{N}$ such that $\operatorname{supp} \varphi \subseteq (-\infty, t_N)$, hence $S_{i+1}(\varphi) - S_i(\varphi) = 0$ for all $i \geq N$, or equivalently,

$$S_i(\varphi) = S_N(\varphi).$$

This implies that the sequence $(S_i(\varphi)) \in (\mathbb{R}^{n \times m})^{\mathbb{N}}$ converges for every test function $\varphi \in \mathcal{C}_{\text{pw}}^\infty$, invoking Proposition 2.1.8 yields the assertion. $\boxed{\text{qed}}$

Corollary 3.3.6 (Existence of Φ_{t_0} and Ψ_{t_0})
Let $t_0 \in \mathbb{R}$, $A \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^{n \times n}$, $n \in \mathbb{N}$ and assume that $A[\cdot]_{(-\infty, t_0)} = 0$. Let $\phi(\cdot, t_0) \in (\mathcal{C}_{\text{pw}}^\infty)^{n \times n}$ be the fundamental solution of the classical ODE $\dot{x} = A^{\text{reg}} x$ as in Lemma 3.3.2. Define

$$\begin{aligned}
&\Phi_{t_0,0} := \phi(\cdot, t_0)_{\mathbb{D}}, \\
\forall i \in \mathbb{N}: \quad &\Phi_{t_0,i+1} := \phi(\cdot, t_0)_{\mathbb{D}} + \phi(\cdot, t_0) \int_{t_0} \phi(\cdot, t_0)^{-1} A[\cdot] \Phi_{t_0,i},
\end{aligned}$$

and, for $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$,

$$\Psi_{t_0,0}(f) := \phi(\cdot,t_0)\int_{t_0} \phi(\cdot,t_0)^{-1} f,$$
$$\forall i \in \mathbb{N}: \quad \Psi_{t_0,i+1}(f) := \phi(\cdot,t_0)\int_{t_0} \phi(\cdot,t_0)^{-1}\big(f + A[\cdot]\Psi_{t_0,i}(f)\big).$$

Then

$$\Phi_{t_0} := \lim_{i\to\infty} \Phi_{t_0,i} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$$

and

$$\Psi_{t_0} : (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \to (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n : f \mapsto \lim_{i\to\infty} \Psi_{t_0,i}(f)$$

are well defined. □

Having established the existence of the above limits it remains to show that all solutions of (3.3.1) can be expressed as $x = \Phi_{t_0}x_0 + \Psi_{t_0}(f)$ for some $x_0 \in \mathbb{R}^n$.

Theorem 3.3.7 (Solution formula for distributional ODE)
Consider the distributional ODE (3.3.1) in standard form and assume there exists $t_0 \in \mathbb{R}$ with $A[\cdot]_{(-\infty,t_0)} = 0$. Let the matrix $\Phi_{t_0} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ and the linear operator $\Psi_{t_0} : (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \to (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be given as in Corollary 3.3.6. Then $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is a solution of (3.3.1) if, and only if, there exists $x_0 \in \mathbb{R}^n$ with

$$x = \Phi_{t_0}x_0 + \Psi_{t_0}(f). \tag{3.3.5}$$

Furthermore, the matrix Φ_{t_0} and the operator Ψ_{t_0} have the following properties:

(i) $\Phi_{t_0}(t_0-) = I$,

(ii) $\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\Phi_{t_0} = A\Phi_{t_0}$,

(iii) $\Psi_{t_0}(f)(t_0-) = 0$,

(iv) $\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\big(\Psi_{t_0}(f)\big) = A\Psi_{t_0}(f) + f$.

In particular, x given by (3.3.5) is the only solution with the property $x(t_0-) = x_0$. $\square$

Proof. Step 1: The properties (i)-(iv) of Φ_{t_0} and Ψ_{t_0} are shown. Note that then any x given by (3.3.5) fulfills $x(t_0-) = x_0$.

(i) Since $A[\cdot]_{(-\infty,t_0)} = 0$, it follows from the definition of Φ_{t_0} in Corollary 3.3.6 that $(\Phi_{t_0})_{(-\infty,t_0)} = (\phi(\cdot,t_0)_{\mathbb{D}})_{(-\infty,t_0)}$, where $\phi(\cdot,t_0)$ is the fundamental solution of the classical ODE $\dot{x} = A^{\text{reg}}x$ as in Lemma 3.3.2. Hence $\Phi_{t_0}(t_0-) = \phi(t_0,t_0) = I$.

(ii) For $i \in \mathbb{N}$ and $\Phi_{t_0,i}$ as in Corollary 3.3.6 the product rule (M3) of the Fuchssteiner multiplication together with $\phi(\cdot,t_0)' = A^{\text{reg}}\phi(\cdot,t_0)$ almost everywhere yields

$$\begin{aligned}
\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\Phi_{t_0,i+1} &= (\phi(\cdot,t_0)_{\mathbb{D}})' + \left(\phi(\cdot,t_0)\int_{t_0}\phi(\cdot,t_0)^{-1}A[\cdot]\Phi_{t_0,i}\right)' \\
&= A_{\text{reg}}\phi(\cdot,t_0)_{\mathbb{D}} + A_{\text{reg}}\phi(\cdot,t_0)_{\mathbb{D}}\int_{t_0}\phi(\cdot,t_0)^{-1}A[\cdot]\Phi_{t_0,i} \\
&\quad + \phi(\cdot,t_0)\phi(\cdot,t_0)^{-1}A[\cdot]\Phi_{t_0,i} \\
&= A_{\text{reg}}\Phi_{t_0,i+1} + A[\cdot]\Phi_{t_0,i}.
\end{aligned}$$

Taking the limit $i \to \infty$ on both sides and invoking Proposition 2.1.8 together with the property

$$\forall \varphi \in \mathcal{C}_0^\infty\ \exists N \in \mathbb{N}\ \forall i \geq N: \quad \Phi_{t_0,i}(\varphi) = \Phi_{t_0}(\varphi) \tag{3.3.6}$$

gives

$$\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\Phi_{t_0} = A_{\text{reg}}\Phi_{t_0} + A[\cdot]\Phi_{t_0} = A\Phi_{t_0}.$$

(iii) By definition, $\left(\int_{t_0} D\right)(t_0-) = 0$ for every piecewise-smooth distribution $D \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}$. Multiplication from the left with a piecewise-smooth function does not change this property, hence, for all $i \in \mathbb{N}$, $f \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ and $\Psi_{t_0,i}(f)$ as in Corollary 3.3.6,

$$\Psi_{t_0,i}(f)(t_0-) = 0,$$

furthermore, by $A[\cdot]_{(-\infty,t_0)} = 0$,

$$\big(\Psi_{t_0}(f)\big)_{(-\infty,t_0)} = \big(\Psi_{t_0,i}(f)\big)_{(-\infty,t_0)},$$

hence

$$\Psi_{t_0}(f)(t_0-) = \Psi_{t_0,i}(f)(t_0-) = 0.$$

(iv) For $\Psi_{t_0,i}(f)$ as in Corollary 3.3.6, $f \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ and $i \in \mathbb{N}$, the product rule (M3) of the Fuchssteiner multiplication and $\phi(\cdot,t_0)' = A_{\mathrm{reg}}\phi(\cdot,t_0)$ almost everywhere yields

$$\begin{aligned}
\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\big(\Psi_{t_0,i+1}(f)\big) &= \Big(\phi(\cdot,t_0)_\mathbb{D}\int_{t_0}\phi(\cdot,t_0)^{-1}\big(f + A[\cdot]\Psi_{t_0,i}(f)\big)\Big)' \\
&= A_{\mathrm{reg}}\phi(\cdot,t_0)_\mathbb{D}\int_{t_0}\phi(\cdot,t_0)^{-1}\big(f + A[\cdot]\Psi_{t_0,i}(f)\big) \\
&\quad + \phi(\cdot,t_0)\phi(\cdot,t_0)^{-1}\big(f + A[\cdot]\Psi_{t_0,i}(f)\big) \\
&= A_{\mathrm{reg}}\Psi_{t_0,i+1}(f) + f + A[\cdot]\Psi_{t_0,i}(f).
\end{aligned}$$

Taking the limit $i \to \infty$ on both sides and invoking Proposition 2.1.8 together with the analogon of (3.3.6) yields

$$\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\big(\Psi_{t_0}(f)\big) = A_{\mathrm{reg}}\Psi_{t_0}(f) + f + A[\cdot]\Psi_{t_0}(f) = A\Psi_{t_0}(f) + f.$$

Step 2: It is shown that any x given by (3.3.5) *is a solution of* (3.3.1). By properties (ii) and (iv) it follows that

$$\dot{x} = \big(\Phi_{t_0}x_0 + \Psi_{t_0}(f)\big)' = A\Phi_{t_0}x_0 + A\Psi_{t_0}(f) + f = Ax + f,$$

i.e. x is a solution of (3.3.1).

Step 3: It is shown that any solution of (3.3.1) *is given by* (3.3.5). Let $\xi \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be any solution of (3.3.1) and let $x = \Phi_{t_0}\xi(t_0-) + \Psi_{t_0}(f)$. It must be shown that $\xi = x$ or, equivalently, $e := \xi - x = 0$. It is

$$\dot{e} = \dot{\xi} - \dot{x} = A\xi + f - (Ax + f) = Ae$$

and $e(t_0-) = \xi(t_0-) - x(t_0-) = 0$. Now Lemma 3.3.3 shows that

$$\dot{e} = Ae, \quad e(t_0-) = 0,$$

only has the trivial solution, hence the claim is shown. $\boxed{\text{qed}}$

3.3.2 ITP solutions of distributional ODEs

Theorem 3.3.8 (DAE-regularity of distributional ODEs)
Every distributional ODE is DAE-regular. □

Proof. By Proposition 3.1.5 and Corollary 3.2.3 it suffices to consider a distributional ODE $(E, A) \in \Sigma^{n\times n}$ in standard form and the impulse free case, i.e. $E = I$ and $A[\cdot] = 0$.

Let $\phi(\cdot, t_0) : \mathbb{R} \to \mathbb{R}^{n\times n}$, $t_0 \in \mathbb{R}$, be the fundamental solution of the classical ODE $\dot{x} = A^{\text{reg}} x$ as in Lemma 3.3.2. It will be shown now that the ITP

$$\begin{aligned} x_{(-\infty,t_0)} &= x^0_{(-\infty,t_0)}, \\ \dot{x}_{[t_0,\infty)} &= (Ax + f)_{[t_0,\infty)}, \end{aligned}$$

where $x^0, f \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ and $t_0 \in \mathbb{R}$, has the unique solution

$$x = x^0_{(-\infty,t_0)} + \left(\phi(\cdot, t_0)_{\mathbb{D}} x^0(t_0-) + \phi(\cdot, t_0)_{\mathbb{D}} \int_{t_0} \phi(\cdot, t_0)^{-1}_{\mathbb{D}} f\right)_{[t_0,\infty)}.$$

It must first be shown, that $\dot{x}_{[t_0,\infty)} = (Ax)_{[t_0,\infty)}$. Corollary 2.3.5, Proposition 2.4.8 and $\phi(\cdot, t_0)'_{\mathbb{D}} = A\phi(\cdot, t_0)_{\mathbb{D}}$ yield

$$\begin{aligned} \dot{x}_{[t_0,\infty)} &= -x^0(t_0-)\delta_{t_0} + \left(\phi(\cdot, t_0)'_{\mathbb{D}} x^0(t_0-) + \left(\phi(\cdot, t_0)_{\mathbb{D}} \int_{t_0} \phi(\cdot, t_0)^{-1}_{\mathbb{D}} f\right)'\right)_{[t_0,\infty)} \\ &\quad + \left(\underbrace{\phi(\cdot, t_0)_{\mathbb{D}}(t_0-)}_{=I} x^0(t_0-) + \underbrace{\left(\phi(\cdot, t_0)_{\mathbb{D}} \int_{t_0} \phi(\cdot, t_0)^{-1}_{\mathbb{D}} f\right)(t_0-)}_{=0}\right)\delta_{t_0} \\ &= \left(A\phi(\cdot, t_0)_{\mathbb{D}} x^0(t_0-) + A\phi(\cdot, t_0)_{\mathbb{D}} \int_{t_0} \phi(\cdot, t_0)^{-1}_{\mathbb{D}} f + f\right)_{[t_0,\infty)} \end{aligned}$$

and, since $A[t_0] = 0$,

$$(Ax+f)_{[t_0,\infty)} = \underbrace{\left(Ax^0_{(-\infty,t_0)}\right)_{[t_0,\infty)}}_{=0} \\ + \left(A\phi(\cdot,t_0)_{\mathbb{D}}x^0(t_0-) + A\phi(\cdot,t_0)_{\mathbb{D}}\int_{t_0}\phi(\cdot,t_0)^{-1}_{\mathbb{D}}f + f\right)_{[t_0,\infty)},$$

which shows that x is a solution of the ITP.

It remains to show that the proposed solution is unique. Assume that $x_1, x_2 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ are solutions of the same ITP, then $e = x_1 - x_2$ fulfills $e_{(-\infty,t_0)} = 0$ and

$$\dot{e}_{[t_0,\infty)} = (Ae)_{[t_0,\infty)}.$$

Note that $\dot{e}_{(-\infty,t_0)} = 0 = (Ae)_{(-\infty,t_0)}$, hence e is a solution of $\dot{e} = Ae$ with $e(t_0-) = 0$ and Lemma 3.3.3 shows that $e = 0$, hence x as given above is the only solution of the ITP. $\boxed{\text{qed}}$

The result of Theorem 3.3.8 is only of qualitative nature it is not obvious how the solution of an ITP looks like, although all theoretical results are already available. The following corollary summarizes all the previous results and gives an explicit formula for an ITP solution of a distributional ODE.

Corollary 3.3.9 (ITP solution of a distributional ODE)
Let $(E,A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$, be a distributional ODE and consider the corresponding ITP with initial time $t_0 \in \mathbb{R}$, initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ and inhomogeneity $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$. Let $\widehat{A} := (E^{-1}A)_{(t_0,\infty)}$ and choose $\widehat{\Phi}_{t_0}$ and $\widehat{\Psi}_{t_0}$ for the distributional ODE $\dot{x} = \widehat{A}x$ as in Theorem 3.3.7. Then the unique solution of the ITP is given by

$$x = x^0_{(-\infty,t_0)} + \left(\widehat{\Phi}_{t_0}x^0(t_0-) + \widehat{\Psi}_{t_0}\left(E^{-1}\left(f_{[t_0,\infty)} + A[t_0]x^0 - E[t_0]\dot{x}^0\right)\right)\right)_{[t_0,\infty)}.$$

$\square$

3.3.3 On the dimension of the solution space of distributional ODEs

For classical ODEs $\dot{x} = Ax$ with some (measurable) matrix function $A : \mathbb{R} \to \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, it is well known that the solution space of $\dot{x} = Ax$ is isomorphic to $\mathbb{R}^n$, in particular its dimension is n. Theorem 3.3.7 shows that this property remains true provided there exists $t_0 \in \mathbb{R}$ such that $A[\cdot]_{(-\infty,t_0)} = 0$ in (3.3.1), i.e. the following corollary holds:

Corollary 3.3.10 (Dimension of solution space for impulse-free past)
Consider a distributional ODE $(E, A) \in \Sigma^{n\times n}$ and assume there exists $t_0 \in \mathbb{R}$ such that $(E^{-1}A)[\cdot]_{(-\infty,t_0)} = 0$. Then

$$\dim\{\ x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \ \mid E\dot{x} = Ax\ \} = n. \qquad \square$$

The following example shows that the condition $(E^{-1}A)[\cdot]_{(-\infty,t_0)} = 0$ for some $t_0 \in \mathbb{R}$ is important in the above result.

Example 3.3.11 (Distributional ODE with trivial solution space)
Consider the distributional ODE

$$\dot{x} = -\sum_{i\in\mathbb{Z}} \delta_i x.$$

First observe that $\dot{x}_{(i,i+1)} = 0$ for all $i \in \mathbb{Z}$, hence, with the same argument as in the proof of Lemma 3.3.3, it follows that x must be constant on $(i, i+1)$, $i \in \mathbb{Z}$. Evaluating the impulsive term at $i \in \mathbb{Z}$ yields $\big(x(i+) - x(i-)\big)\delta_i + \big(x[i]\big)' = x(i-)\delta_i$, hence $x(i+) =$ and $x[i] = 0$ for all $i \in \mathbb{R}$. Altogether this implies that only the trivial solution $x = 0$ solves the above distributional ODE, hence

$$\dim\{\ x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \ \mid E\dot{x} = Ax\ \} = 0. \qquad \square$$

For the space of consistent solutions there is at least an upper bound for the dimension as the following proposition shows.

Proposition 3.3.12 (Upper bound for solution space)
Every distributional ODE $(E, A) \in \Sigma^{n\times n}$ satisfies

$$\dim\{\ x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \ \mid E\dot{x} = Ax\ \} \leq n. \qquad \square$$

Proof. Without restriction assume $E = I$. For a given coefficient matrix A, Corollary 3.3.9 reveals that the solutions on $[t_0, \infty)$, $t_0 \in \mathbb{R}$, only depend on $x^{(i)}(t_0-)$, $i \in \mathbb{N}$. Observe that

$$\begin{aligned}
\dot{x}(t_0-) &= A(t_0-)x(t_0-)\\
\ddot{x}(t_0-) &= (Ax)'(t_0-) = A'(t_0-)x(t_0-) + A(t_0-)\dot{x}(t_0-)\\
&= (A' + A^2)(t_0-)x(t_0-)\\
&\ \ \vdots\\
x^{(i)}(t_0-) &= (Ax)^{(i-1)}(t_0-) = \sum_{j=0}^{i-1} \tbinom{i-1}{j} A^{(i-1-j)}(t_0-)x^{(j)}(t_0-),
\end{aligned}$$

hence, inductively, all derivatives $x^{(i)}(t_0-)$, $i \in \mathbb{N}$, are determined by $x(t_0-)$. Altogether this implies that for two consistent solutions x^1, x^2 of the distributional ODE (3.3.1) the following implication holds for all $t_0 \in \mathbb{R}$:

$$x^1(t_0-) = x^2(t_0-) \quad \Rightarrow \quad x^1_{[t_0,\infty)} = x^2_{[t_0,\infty)}. \tag{3.3.7}$$

Let $x_1, x_2, \ldots, x_k \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$, $k \in \mathbb{N}$, be linearly independent solutions of (3.3.1) and let $X := [x_1, x_2, \ldots, x_k] \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times k}$. It will be shown that there exist a $t_0 \in \mathbb{R}$ such that $X(t_0-)$ has full column rank and, therefore, $k \leq n$ holds.

For $t \in \mathbb{R}$ let $\mathcal{K}_t := \ker X(t-) \subseteq \mathbb{R}^k$, then, by (3.3.7) and linearity,

$$\forall \alpha \in \mathcal{K}_t : \quad X_{[t,\infty)}\alpha = 0,$$

hence

$$\forall s \geq t : \quad \mathcal{K}_s \supseteq \mathcal{K}_t.$$

Let $\mathcal{K} := \bigcap_{t\in\mathbb{R}} \mathcal{K}_t$ and, seeking a contradiction, assume $\mathcal{K} \neq \{0\}$. For $\alpha \in \mathcal{K} \setminus \{0\}$ it holds that

$$\forall t \in \mathbb{R} : \quad X_{[t,\infty)}\alpha = 0,$$

hence $X\alpha = 0$ which contradicts the linear independence of x_1, x_2, …, x_k. Therefore, $\mathcal{K} = \{0\}$. This implies that there exists $t_0 \in \mathbb{R}$ such that $\mathcal{K}_{t_0} = \{0\}$ or, equivalently, that $X(t_0-)$ has full column rank. $\boxed{\text{qed}}$

Consider now ITP solutions for a distributional ODE $(E, A) \in \Sigma^{n\times n}$, then a direct consequence from Corollary 3.3.9 is the following result.

Corollary 3.3.13 (Dimension of solution space, impulse free case)
Let $(E, A) \in \Sigma^{n\times n}$ be a distributional ODE with $E[t_0] = 0 = A[t_0]$ for $t_0 \in \mathbb{R}$. Then

$$\dim \left\{ x_{[t_0,\infty)} \;\middle|\; (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)} \right\} = n \qquad \square$$

The following examples show what could happen when the assumption $E[t_0] = 0 = A[t_0]$ does not hold.

Examples 3.3.14 (Dimension of ITP solution space)
In the following, distributional ODEs $(E, A) \in \Sigma^{n\times n}$ and the corresponding ITPs with initial time t_0 and initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ are considered.

(i) Let $E = I$ and $A = -\delta_0 I$, then all solutions $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of the ITP fulfill $x_{[t_0,\infty)} = 0$, hence

$$\dim \left\{ x_{[t_0,\infty)} \;\middle|\; (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)} \right\} = 0.$$

(ii) Let $E = I + \delta_0 I$ and $A = 0$, then all solutions $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of the ITP fulfill

$$x_{[t_0,\infty)} = \big(x^0(t_0-) - \dot{x}^0(t_0-)\big)\mathbb{1}_{[t_0,\infty)_{\mathbb{D}}}$$

hence the solution depends also on the derivative of the initial trajectory, however the dimension of the solution space is not $2n$ but remains n:

$$\dim \left\{ x_{[t_0,\infty)} \;\middle|\; (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)} \right\} = n.$$

(iii) Let $E = I$, $A = \delta_0' I$, then all solutions $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of the ITP fulfill

$$x_{[t_0,\infty)} = \big(x^0(t_0-) - \dot{x}^0(t_0-)\big)\mathbb{1}_{[t_0,\infty)_{\mathbb{D}}} + x(t_0-)\delta_t,$$

again the solution depends on the derivative of the initial trajectory but this time this increases the dimension of the solution space:

$$\dim\left\{\ x_{[t_0,\infty)}\ \middle|\ (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)}\ \right\} = 2n.$$

Note that the dimension of the solution space can be arbitrarily high, just replace δ_0' above by $\delta_0^{(k)}$ for some $k \in \mathbb{N}$. ◻

There is a similar result as Proposition 3.3.12 for ITP solutions, but it only holds when one considers the solution in the "strict" future, i.e. the restriction of the solution to the open interval (t_0, ∞).

Proposition 3.3.15 (Upper bound for dimension)
Consider a distributional ODE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$, with initial time $t_0 \in \mathbb{R}$, then

$$\dim\left\{\ x_{(t_0,\infty)}\ \middle|\ (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)}\ \right\} \le n.$$ ◻

Proof. From (3.3.4) together with the regularity of a distributional ODE, it follows that $x_{(t_0,\infty)}$ is uniquely given by $x(t_0+)$, hence the dimension of the solution space for $x_{(t_0,\infty)}$ fulfills

$$\begin{aligned}&\dim\left\{\ x_{(t_0,\infty)}\ \middle|\ (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)}\ \right\}\\ &\quad = \dim\left\{\ x_0 \in \mathbb{R}^n\ \middle|\ \exists \text{ ITP solution } x \text{ with } x(t_0+) = x_0\ \right\} \le n.\end{aligned}$$ qed

3.4 Pure distributional DAEs

Definition 3.4.1 (Pure distributional DAE)
A DAE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$, is called a *pure distributional DAE* if, and only if, A is invertible and $(A^{-1}E)_{\text{reg}}$ is a strictly lower triangular matrix. A pure distributional DAE is in *standard form* if, and only if, $A = I$. ◻

As in Section 3.3.1, it suffices to consider the pure distributional DAE in standard form

$$N\dot{x} = x + f, \tag{3.4.1}$$

where $N \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ is such that N_{reg} is a strictly lower triangular matrix. The key "ingredient" of the solution theory for pure distributional DAEs is the following lemma.

Lemma 3.4.2 (Nilpotency of $N\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}$)
Let $N \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$, $n \in \mathbb{N}$, and assume that N_{reg} is a strictly lower triangular matrix. Consider the linear operator

$$N\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t} : (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \to (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n, \quad x \mapsto \left(N\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\right)(x) = N\dot{x}$$

with its corresponding functional power $\left(N\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\right)^i : (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \to (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$, $i \in \mathbb{N}$. Then $N\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}$ is nilpotent, i.e. there exists $\nu \in \mathbb{N}$ such that $(N\frac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t})^\nu = 0$. □

Proof. In the following a square matrix $M \in \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, is called *lower triangular of order* $i \in \mathbb{N}$ if, and only if, $M_{kl} = 0$ for all $k, l \in \mathbb{N}$ with $k < l + i$. Note that any lower triangular matrix is a lower triangular matrix of order zero and a lower triangular matrix of order $i \geq n$ is the zero matrix. Furthermore, the product of two lower triangular matrices M_1, M_2 of orders $i_1, i_2 \in \mathbb{N}$, resp., is a lower triangular matrix of order $i_1 + i_2$.

To prove the Lemma, it is shown inductively that, for $\nu \in \mathbb{N}$,

$$\left(N\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\right)^\nu = \sum_{i=0}^{\nu} \left(M_{\nu,i} + \sum_{k=0}^{\nu-1} \sum_{j=1}^{m_{\nu,i,k}} P_{\nu,i,k,j} H_{\nu,i,k,j} Q_{\nu,i,\nu-1-k,j} \right) (\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t})^i,$$

where $m_{\nu,i,k} \in \mathbb{N}$, $M_{\nu,i} \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}^{n\times n}$ is a lower triangular matrix of order ν. The matrices $P_{\nu,i,k,j} \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}^{n\times n}$ and $Q_{\nu,i,\nu-1-k,j} \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}^{n\times n}$ are impulse-free and lower triangular matrices of order k and $\nu-1-k$, resp., finally, $H_{\nu,i,k,j} = H_{\nu,i,k,j}[\cdot] \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}^{n\times n}$ is a purely impulsive matrix.

Since

$$\left(N\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\right)^0 = \mathrm{id} = I\left(\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\right)^0,$$

and $I \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}^{n\times n}$ is a lower triangular matrix of order 0 the claim is shown for $\nu = 0$ with $M_{0,0} = I$. Now assume that the claim holds for some $\nu \in \mathbb{N}$, then

$$\begin{aligned}
\left(N\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\right)^{\nu+1} &= (N_{\mathrm{reg}} + N[\cdot])\left(\left(N\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\right)^{\nu}\right)' \\
&= \sum_{i=0}^{\nu}\Bigg((N_{\mathrm{reg}} + N[\cdot])M'_{\nu,i} + \\
&\qquad + \sum_{k=0}^{\nu-1}\sum_{j=1}^{m_{\nu,i,k}} N_{\mathrm{reg}}(P_{\nu,i,k,j})' H_{\nu,i,k,j} Q_{\nu,i,\nu-1-k,j} \\
&\qquad + \sum_{k=0}^{\nu-1}\sum_{j=1}^{m_{\nu,i,k}} N_{\mathrm{reg}} P_{\nu,i,k,j}\big(H_{\nu,i,k,j}\big)' Q_{\nu,i,\nu-1-k,j} \\
&\qquad + \sum_{k=0}^{\nu-1}\sum_{j=1}^{m_{\nu,i,k}} N_{\mathrm{reg}} P_{\nu,i,k,j} H_{\nu,i,k,j} (Q_{\nu,i,\nu-1-k,j})'\Bigg)\left(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\right)^{i} \\
&\quad + \sum_{i=0}^{\nu}\Bigg((N_{\mathrm{reg}} + N[\cdot])M_{\nu,i} + \\
&\qquad + \sum_{k=0}^{\nu-1}\sum_{j=1}^{m_{\nu,i,k}} N_{\mathrm{reg}} P_{\nu,i,k,j} H_{\nu,i,k,j} Q_{\nu,i,\nu-1-k,j}\Bigg)\left(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\right)^{i+1} \\
&= \sum_{i=0}^{\nu+1}\Bigg(M_{\nu+1,i} + \\
&\qquad + \sum_{k=0}^{\nu}\sum_{j=1}^{m_{\nu+1,i,k}} P_{\nu+1,i,k,j} H_{\nu+1,i,k,j} Q_{\nu+1,i,\nu-k,j}\Bigg)\left(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\right)^{i},
\end{aligned}$$

where

$$\begin{aligned}
M_{\nu+1,0} &= N_{\mathrm{reg}} M'_{\nu,0}, \\
M_{\nu+1,i} &= N_{\mathrm{reg}} M'_{\nu,i} + N_{\mathrm{reg}} M_{\nu,i\text{-}1}, \qquad i = 1,\dots,\nu \\
M_{\nu+1,\nu+1} &= N_{\mathrm{reg}} M_{\nu,\nu},
\end{aligned}$$

$$\begin{aligned}
P_{\nu+1,i,0,1} &= I, && i = 0,\dots,\nu+1\\
H_{\nu+1,i,0,1} &= N[\cdot], && i = 0,\dots,\nu+1\\
Q_{\nu+1,0,\nu,1} &= (M'_{\nu,i})_{\text{reg}}, &&\\
Q_{\nu+1,i,\nu,1} &= (M'_{\nu,i} + M_{\nu,i\text{-}1})_{\text{reg}} && i = 1,\dots,\nu\\
Q_{\nu+1,\nu+1,\nu,1} &= (M_{\nu,\nu})_{\text{reg}}, &&\\
m_{\nu+1,i,0} &= 1, && i = 0,\dots,\nu+1
\end{aligned}$$

and for $k = 1,\dots,\nu$,

$$\begin{aligned}
m_{\nu+1,0,k} &= 3m_{\nu,0,k\text{-}1}, &&\\
m_{\nu+1,i,k} &= 3m_{\nu,i,k\text{-}1} + m_{\nu,i\text{-}1,k\text{-}1}, && i = 1,\dots,\nu\\
m_{\nu+1,\nu+1,k} &= m_{\nu,\nu,k\text{-}1}, &&\\
P_{\nu+1,\nu+1,k,j} &= N_{\text{reg}} P_{\nu,\nu,k\text{-}1,j}, && j = 1,\dots,m_{\nu,\nu,k\text{-}1},\\
H_{\nu+1,\nu+1,k,j} &= H_{\nu,\nu,k\text{-}1,j}, && j = 1,\dots,m_{\nu,\nu,k\text{-}1},\\
Q_{\nu+1,\nu+1,\nu\text{-}k,j} &= Q_{\nu,\nu,\nu\text{-}k,j}, && j = 1,\dots,m_{\nu,\nu,k\text{-}1},
\end{aligned}$$

and additionally for $i = 0,\dots,\nu$

$$\begin{aligned}
P_{\nu+1,i,k,3j\text{-}2} &= N_{\text{reg}} (P'_{\nu,i,k\text{-}1,j})_{\text{reg}}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
H_{\nu+1,i,k,3j\text{-}2} &= H_{\nu,i,k\text{-}1,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
Q_{\nu+1,i,\nu\text{-}k,3j\text{-}2} &= Q_{\nu,i,\nu\text{-}k,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
P_{\nu+1,i,k,3j\text{-}1} &= N_{\text{reg}} P_{\nu,i,k\text{-}1,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
H_{\nu+1,i,k,3j\text{-}1} &= H'_{\nu,i,k\text{-}1,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
Q_{\nu+1,i,\nu\text{-}k,3j\text{-}1} &= Q_{\nu,i,\nu\text{-}k,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
P_{\nu+1,i,k,3j} &= N_{\text{reg}} P_{\nu,i,k\text{-}1,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
H_{\nu+1,i,k,3j} &= H_{\nu,i,k\text{-}1,j}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
Q_{\nu+1,i,\nu\text{-}k,3j} &= (Q'_{\nu,i,\nu\text{-}k,j})_{\text{reg}}, && j = 1,\dots,m_{\nu,i,k\text{-}1},\\
P_{\nu+1,i,k,j+3m_{\nu,i,k\text{-}1}} &= N_{\text{reg}} P_{\nu,i\text{-}1,k\text{-}1,j}, && i \neq 0,\ j = 1,\dots,m_{\nu,i\text{-}1,k\text{-}1},\\
H_{\nu+1,i,k,j+3m_{\nu,i,k\text{-}1}} &= H_{\nu,i\text{-}1,k\text{-}1,j}, && i \neq 0,\ j = 1,\dots,m_{\nu,i\text{-}1,k\text{-}1},\\
Q_{\nu+1,i,\nu\text{-}k,j+3m_{\nu,i,k\text{-}1}} &= Q_{\nu,i\text{-}1,\nu\text{-}k,j}, && i \neq 0,\ j = 1,\dots,m_{\nu,i\text{-}1,k\text{-}1}.
\end{aligned}$$

This proves the claim. For $\nu \geq n$ it is $M_{\nu,i} = 0$ for all $i = 0, \ldots, \nu$, furthermore for $\nu \geq 2n$ either $k \geq n$ or $\nu - 1 - k \geq n$, hence either $P_{\nu,i,k,j} = 0$ or $Q_{\nu,i,\nu-1-k,j} = 0$. Altogether this yields for $\nu \geq 2n$

$$(N\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{\nu} = 0.$$

qed

It is now very simple to characterize all solutions of the pure DAE (3.4.1): the pure DAE in standard form (3.4.1) can be rewritten as a linear operator equation

$$(N\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} - I)(x) = f,$$

and since by Lemma 3.4.2 the operator $N\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}$ is nilpotent it is easy to see that the operator $(N\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} - I) : (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \to (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is bijective with inverse operator

$$(N\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} - I)^{-1} = -\sum_{i=0}^{\nu-1}(N\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{i},$$

where $\nu \in \mathbb{N}$ is such that $(N\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{\nu} = 0$. This already yields the following theorem.

Theorem 3.4.3 (Unique solution of pure distributional DAE)
Every pure distributional DAE $(E, A) \in \Sigma^{n\times n}$ with inhomogeneity $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is uniquely solvable and its solution is given by

$$x = -\sum_{i=0}^{\nu-1}(A^{-1}E\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{i}(A^{-1}f),$$

where $\nu \in \mathbb{N}$ is such that $(A^{-1}E\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{\nu} = 0$. In particular, the homogeneous pure distributional DAE has only the trivial solution. □

Proof. Clearly, x solves $E\dot{x} = Ax + f$ if, and only if, x solves $A^{-1}E\dot{x} = x + A^{-1}f$. By Definition $(A^{-1}E)_{\mathrm{reg}}$ is a strictly lower triangular matrix and Lemma 3.4.2 yields the above mentioned invertibility of the operator $(A^{-1}E\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} - I)$. qed

Corollary 3.4.4 (DAE-regularity of pure distributional DAEs)
Every pure distributional DAE is DAE-regular. In particular, for any pure distributional DAE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$, with inhomogeneity $f \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ the corresponding ITP with initial time $t_0 \in \mathbb{R}$ and initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ has the unique solution

$$x = -\sum_{i=0}^{\nu-1} \big((A^{-1}E)_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}\big)^i \big(f_{[t_0,\infty)} - x^0_{(-\infty,t_0)} + A^{-1}[t_0](E\dot{x}^0 - Ax^0)\big),$$

where $\nu \in \mathbb{N}$ is such that $(A^{-1}E)^\nu = 0$. □

Proof. Consider first a pure DAE in standard form (3.4.1) with initial time $t_0 \in \mathbb{R}$ and initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$. By Theorem 3.1.7 (N, I) is DAE-regular if, and only if,

$$N_{\mathrm{itp}}\dot{x} = x + f_{\mathrm{itp}},$$

is uniquely solvable, where $N_{\mathrm{itp}} = N_{[t_0,\infty)}$ and $f_{\mathrm{itp}} = x^0_{(-\infty,t_0)} + f_{[t_0,\infty)}$. Since the latter equation is itself a pure distributional DAE, Theorem 3.4.3 shows that there exists a unique solution. Proposition 3.1.5 yields that also general pure distributional DAEs are regular and simple calculations in the spirit of Remark 3.1.6 yield the explicit solution formula. qed

This section will be concluded with a result on the dimension of the solution spaces of homogeneous pure distributional DAEs, i.e. pure distributional DAEs (3.4.1) for which the inhomogeneity f is zero.

Proposition 3.4.5 (Dimension of solutions space)
Consider a pure distributional DAE $(E, A) \in \Sigma^{n\times n}$.

(i) $\dim \{\ x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \ |\ E\dot{x} = Ax\ \} = 0$.

(ii) $\dim \big\{\ x_{(t_0\infty)} \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \ \big|\ (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)}\ \big\} = 0$ for any $t_0 \in \mathbb{R}$.

(iii) If $E[t_0] = 0 = A[t_0]$ for some $t_0 \in \mathbb{R}$, then

$$\dim \big\{\ x_{[t_0,\infty)} \ \big|\ (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)}\ \big\} \leq \dim \operatorname{im} \big(A^{-1}E\big)(t_0+).$$ □

Proof. The first two dimension formulae follow directly from the explicit solution formulae given in Theorem 3.4.3 and Corollary 3.4.4.

To prove (iii), let $N := A^{-1}E$. Then the solution of the ITP with initial time $t_0 \in \mathbb{R}$ and initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is given by

$$x = \sum_{i=0}^{\nu-1} (N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^i x^0_{(-\infty,t_0)}.$$

The assumption $E[t_0] = 0 = A[t_0]$ implies that $N[t_0] = 0$, hence

$$\begin{aligned} N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} x^0_{(-\infty,t_0)} &= N_{[t_0,\infty)}(\dot{x}^0_{(-\infty,t_0)} - x^0(t_0-)\delta_{t_0}) \\ &= -N(t_0+)x^0(t_0-)\delta_{t_0} \end{aligned}$$

and it easily follows that

$$\sum_{i=2}^{\nu-1} (N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^{i-1} N(t_0+)x^0(t_0-)\delta_{t_0} = M_{t_0} N(t_0+)x^0(t_0-),$$

for some distributional matrix $M_{t_0} \in \mathbb{D}_{\{t_0\}}$ with point support. Therefore,

$$x_{[t_0,\infty)} = (I + M_{t_0})N(t_0+)x^0(t_0-)$$

and the claim follows. qed

Remark 3.4.6 (Exact dimension of solution space)
Let $(E, A) \in \Sigma^{n\times n}$ be a pure distributional DAE with $E[\cdot] = 0 = A[\cdot]$ and let $N := A^{-1}E$ and choose $\nu \in \mathbb{R}$ such that $(N \frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^\nu = 0$. It is possible (but also very technical) to calculate matrices $M_{t_0,j} \in \mathbb{R}^{n\times n}$, $j = 0, \ldots, \nu-1$, which only depend on $N^{(i)}(t_0+)$, $i = 0, \ldots, \nu-1$, such that

$$x_{[t_0,\infty)} = \sum_{i=0}^{\nu-1} (N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^i x^0_{(-\infty,t_0)} = \sum_{j=0}^{\nu-1} M_{t_0,j}\, x(t_0-)\delta_0.$$

The dimension of the solution space is then exactly given by

$$\dim \left\{ \, x_{[t_0,\infty)} \;\middle|\; (E\dot{x})_{[t_0,\infty)} = (Ax)_{[t_0,\infty)} \, \right\} = \max_j \dim \operatorname{im} M_{t_0,j}. \qquad \square$$

As an illustration of the above results consider the following examples, in particular it is shown that if the assumption $A[t_0] = 0 = E[t_0]$ does not hold, then the dimension of the solution space as in Proposition 3.4.5(iii) can be arbitrarily large.

Example 3.4.7 (ITP solution space for pure distributional DAEs)
Consider the pure distributional DAE (3.4.1) in standard form with $N \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$ and let $t_0 \in \mathbb{R}$ and $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be some initial time and some initial trajectory.

(i) Let $N \in \mathbb{R}^{n\times n}$ be some constant nilpotent matrix, then it follows that

$$x_{[t_0,\infty)} = \sum_{i=0}^{n-1} N^i x(t_0-)\delta_{t_0}^{(i)},$$

hence the dimension of the solution space for $x_{[t_0,\infty)}$ is exactly $\dim \operatorname{im} N(t_0+) = \dim \operatorname{im} N$.

(ii) Let $n = 1$ and $N = \delta_0^{(k)} \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ for some k, then

$$x_{[t_0,\infty)} = \delta_0^{(k)} \dot{x}^0_{(-\infty,t_0)} = \sum_{i=0}^{k} (-1)^i (x^0)^{(i+1)}(t_0-)\delta_t^{(k-i)},$$

hence, by 2.1.14, the solution space has dimension $k+1$, i.e. the dimension can be arbitrarily large. □

3.5 Generalized Weierstraß form

A direct consequence of Proposition 3.2.2 together with Theorem 3.3.8 and Corollary 3.4.4 is the following result.

Corollary 3.5.1 (Generalized Weierstraß form)
Consider a distributional DAE $(E, A) \in \Sigma^{n\times n}$, $n \in \mathbb{N}$. If there exist invertible matrices $S, T \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$, a locally finite set $\{t_i \in \mathbb{R} \,|\, i \in \mathbb{Z}\}$, a family of matrices $J_i \in (\mathcal{C}^\infty_{\mathrm{pw}})^{n_i\times n_i}$, $i \in \mathbb{Z}$, $0 \le n_i \le n$, and a family of

strictly lower triangular matrices $N_i \in (\mathcal{C}^\infty_{\text{pw}})^{(n-n_i)\times(n-n_i)}$, $i \in \mathbb{Z}$, such that

$$\begin{aligned}\big((SET)_{\text{reg}}, (SAT - SET')_{\text{reg}}\big)\\ = \left(\sum_{i\in\mathbb{Z}} \begin{bmatrix} I & \\ & N_{i\mathbb{D}} \end{bmatrix}_{[t_i,t_{i+1})}, \sum_{i\in\mathbb{Z}} \begin{bmatrix} J_{i\mathbb{D}} & \\ & I \end{bmatrix}_{[t_i,t_{i+1})}\right),\end{aligned} \tag{3.5.1}$$

then (E, A) is DAE-regular. In the following (3.5.1) will be called *generalized Weierstraß form.* □

It is still an open question whether the existence of a generalized Weierstraß form for a distributional DAE $(E, A) \in \Sigma^{n\times n}$ is also a necessary condition for DAE-regularity of (E, A). However, there are special cases where DAE-regularity is equivalent to the existence of a generalized Weierstraß form.

Theorem 3.5.2 (Piecewise-constant coefficients)
Let $(E, A) = \big(\sum_{i\in\mathbb{Z}} E_{i[t_i,t_{i+1})}, \sum_{i\in\mathbb{Z}} A_{i[t_i,t_{i+1})}\big)$ where $\{\ t_i \in \mathbb{R}\ \mid i \in \mathbb{Z}\ \}$ is a locally finite set and each matrix pair $(E_i, A_i) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, $i \in \mathbb{Z}$, is constant. Then the corresponding distributional DAE (E, A) is DAE-regular if, and only if, (E_i, A_i) is regular in the classical sense, i.e. $\det(E_i s - A_i) \in \mathbb{R}[s]\backslash\{0\}$, for all $i \in \mathbb{Z}$. In particular, the corresponding distributional DAE (E, A) is DAE-regular if, and only if, it can be put into the generalized Weierstrass normal form (3.5.1). Furthermore, for two solutions $x, y \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ the following implication holds for all $t_0 \in \mathbb{R}$

$$x(t_0-) = y(t_0-) \quad \Rightarrow \quad x_{[t_0,\infty)} = y_{[t_0,\infty)}$$

□

Proof. It is well known that any matrix pair $(E_i, A_i) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $i \in \mathbb{Z}$, is regular in the classical sense if, and only if, there exists invertible matrices $S_i, T_i \in \mathbb{R}^{n\times n}$ such that (E_i, A_i) is put into the Weierstraß normal form:

$$(S_i E_i T_i, S_i A_i T_i) = \left(\begin{bmatrix} I & 0 \\ 0 & N_i \end{bmatrix}, \begin{bmatrix} J_i & 0 \\ 0 & I \end{bmatrix}\right)$$

where $J_i \in \mathbb{R}^{n_{i,1}\times n_{i,1}}$, $n_{i,1} \in \mathbb{N}$, and $N_i \in \mathbb{R}^{(n-n_{i,1})\times(n-n_{i,1})}$ are matrices in (real) Jordan normal form, N_i is nilpotent and I stands for different identity matrices of appropriate size. Let $S = \sum_{i\in\mathbb{Z}} S_{i[t_i,t_{i+1})}$ and $T = \sum_{i\in\mathbb{Z}} T_{i[t_i,t_{i+1})}$, then clearly $(SET, SAT - SET')_{\text{reg}} = (SET, SAT)$ is in generalized Weierstraß form (3.5.1) and Corollary 3.5.1 implies that the corresponding distributional DAE (E, A) is DAE-regular. If for some $i \in \mathbb{Z}$ the matrix pair (E_i, A_i) is not regular in the classical sense it can be shown, see e.g. [KM06, Thm 2.14], that there exist non-trivial solutions of the homogeneous DAE $E_i\dot{x} = A_i x$ with $x_{(\infty,t_i)} = 0$, hence (E, A) is not uniquely solvable and therefore (E, A) is not DAE-regular.

To show that each solutions x is uniquely defined on $[t_0, \infty)$ by its value $x(t_0-)$ it suffices to consider the single constant coefficient ITP

$$\begin{aligned} (E\dot{x})_{[t_0,\infty)} &= (Ax)_{[t_0,\infty)}, \\ x_{(-\infty,t_0)} &= x^0_{(-\infty,t_0)}, \end{aligned} \tag{3.5.2}$$

where $x^0 \in \mathbb{D}_{\text{pw}\mathcal{C}^\infty}$ is an arbitrary initial trajectory and (E, A) is some regular matrix pair. It will be shown that $x_{[t_0,\infty)}$ only depends on $x(t_0-)$, the claim of Theorem 3.5.2 follows then easily by induction.

Since (E, A) is regular, there exist matrices S and T such that $(SET, SAT) = \left(\left[\begin{smallmatrix} I & 0 \\ 0 & N \end{smallmatrix}\right], \left[\begin{smallmatrix} J & 0 \\ 0 & I \end{smallmatrix}\right]\right)$ for some nilpotent matrix $N \in \mathbb{R}^{n_2\times n_2}$, $n_2 \in \mathbb{N}$, and some matrix $J \in \mathbb{R}^{n_1\times n_1}$, $n_1 = n - n_2$. Let $z = T^{-1}x$ and $z^0 = T^{-1}x^0$, then, clearly, x solves (3.5.2) if, and only if, $z = T^{-1}x$ solves

$$\begin{aligned} \left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \dot{z}\right)_{[t_0,\infty)} &= \left(\begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} z\right)_{[t_0,\infty)}, \\ z_{(-\infty,t_0)} &= z^0_{(-\infty,t_0)}. \end{aligned}$$

The solution z is now given by two independent solutions of a distributional ODE and a pure distributional DAE, since the coefficient matrices are impulse free, the solution formulae from Corollary 3.3.9 and Corollary 3.4.4 reveal that $z_{[t_0,\infty)}$ only depends on $z^0(t_0-)$ which implies that every solution x of the ITP (3.5.2) is uniquely given by $x^0(t_0-)$. $\boxed{\text{qed}}$

In the following, DAEs with real analytical coefficients will be considered. A function $f : \mathbb{R} \to \mathbb{R}$ is called *real analytical* at $t \in \mathbb{R}$ if, and only if, all derivatives $f^{(i)}(t)$, $i \in \mathbb{N}$ at t exists and if there exists $\varepsilon > 0$ such that

$$\forall \tau \in (-\varepsilon, \varepsilon) : \quad f(t+\tau) = \sum_{i\in\mathbb{N}} \frac{f^{(i)}(t)}{i!}\tau^i.$$

A function $f : \mathbb{R} \to \mathbb{R}$ is called real analytical on some interval $H \subseteq \mathbb{R}$ if, and only if f is real analytical at t for all $t \in H$. For later results the following lemma is needed.

Lemma 3.5.3
Let $f : \mathbb{R} \to \mathbb{R}$ be real analytical on some closed interval $[s,t] \subseteq \mathbb{R}$, then there exists a smooth $g \in \mathcal{C}^\infty$ such that $g_{[s,t]} = f_{[s,t]}$. $\square$

Proof. By analyticity of f there exists $\varepsilon > 0$ such that $f(s-\tau) = \sum_{i\in\mathbb{N}} \frac{f^{(i)}(s)}{i!}\tau^i$ and $f(t+\tau) = \sum_{i\in\mathbb{N}} \frac{f^{(i)}(t)}{i!}\tau^i$ for all $0 \leq \tau < \varepsilon$, hence f is smooth on $(s-\varepsilon, t+\varepsilon)$, see e.g. [AE05, Cor. V.3.2]. Choose a smooth function $\mathbb{1}_{[s,t]}^{\varepsilon/2} \in \mathcal{C}^\infty$ such that $\mathbb{1}_{[s,t]}^{\varepsilon/2}(\tau) = 1$ for all $\tau \in [s,t]$ and $\mathbb{1}_{[s,t]}^{\varepsilon/2}(\tau) = 0$ for all $\tau \in \mathbb{R}\backslash(s-\varepsilon/2, t+\varepsilon/2)$. Then $g := \mathbb{1}_{[s,t]}^{\varepsilon/2} f$ is smooth and $g_{[s,t]} = f_{[s,t]}$. $\boxed{\text{qed}}$

In view of [CP83], a matrix pair (E,A) with real analytical coefficients on some closed interval $[s,t] \subseteq \mathbb{R}$ is called *analytically solvable* on the interval $[s,t]$ if and only if for all smooth $f : [s,t] \to \mathcal{C}^\infty$ there exists a classical solution $x : [s,t] \to \mathbb{R}^n$ of $E\dot{x} = Ax + f$ and all solutions are uniquely determined by the value $x(t_0)$ for any fixed $t_0 \in [s,t]$.

In [CP83] it is shown that the matrix pair (E,A) is analytically solvable on $\mathbb{R}$ if, and only if, there exists invertible matrix functions $S, T : \mathbb{R} \to \mathbb{R}^{n\times n}$ which are real analytical on $[s,t]$ such that, on the interval $[s,t]$,

$$(SET, SAT - SET') = \left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right), \tag{3.5.3}$$

where J is some real analytical matrix function and N is a strictly lower triangular real analytical matrix function. With this result the following corollary can be shown easily.

Corollary 3.5.4
Let $(E, A) = \left(\sum_{i\in\mathbb{Z}} E_{i[t_i,t_{i+1})}, \sum_{i\in\mathbb{Z}} A_{i[t_i,t_{i+1})}\right)$ where $\{ t_i \in \mathbb{R} \mid i \in \mathbb{Z} \}$ is a locally finite set and each matrix pair $(E_i, A_i) : \mathbb{R} \to \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, $i \in \mathbb{Z}$, is real analytical and analytically solvable on $[t_i, t_{i+1}]$. Then the corresponding distributional DAE (E, A) is DAE-regular. $\square$

Proof. Since each (E_i, A_i), $i \in \mathbb{Z}$, is analytically solvable there exist invertible matrices S_i, T_i which are analytical on $[t_i, t_{i+1}]$ such that (3.5.3) holds. With $S = \sum_{i\in\mathbb{Z}} S_{i[t_i,t_{i+1})}$ and $T = \sum_{i\in\mathbb{Z}} T_{i[t_i,t_{i+1})}$ it follows that $(SET, SAT - SET')_{\text{reg}}$ is in generalized Weierstraß form (3.5.1). $\boxed{\text{qed}}$

It remains unclear whether DAE-regularity implies analytical solvability for analytical coefficients: for analytical solvability it is assumed that any local solution can be extended to a solution on the whole corresponding interval, but for DAE-regularity it is only assumed that solutions can uniquely be extended into the future. Furthermore, DAE-regularity guaranties the existence of local solution for arbitrary inhomogeneities, but it only guarantees existence of distributional solutions and it is not clear in general when these are actually classical solutions.

4 Switched DAEs

4.1 System class and motivation

In this section switched differential algebraic equations (switched DAEs) of the form

$$E_\sigma \dot{x} = A_\sigma x \tag{4.1.1}$$

will be studied, here $\sigma : \mathbb{R} \to \{1, \ldots, N\}$, $N \in \mathbb{N}$, is a switching signal and $E_p, A_p \in \mathbb{R}^{n \times n}$, $n \in \mathbb{N}$, are constant coefficient matrices for each parameter $p \in \{1, \ldots, N\}$.

Switched DAEs occur, for example, in modeling electrical circuits with switches or when modeling possible faults in systems where each (faulty and non-faulty) configuration is described by a classical DAE $E\dot{x} = Ax$ with constant matrices $E, A \in \mathbb{R}^{n \times n}$.

Throughout this section the following two assumptions will be made.

Assumptions
Consider the switched DAE (4.1.1).

(S1) The switching signal $\sigma : \mathbb{R} \to \{1, 2, \ldots, N\}$ is piecewise-constant with a locally finite set of jump points and right-continuous.

(S2) Each matrix pair (E_p, A_p), $p = 1, \ldots, N$, is regular, i.e. $\det(sE_p - A_p) \in \mathbb{R}[s] \setminus \{0\}$. □

These assumption ensure that the switched DAE (4.1.1) corresponds to a regular distributional DAE as introduced in Section 3.1, i.e. for each initial trajectory there exists a unique distributional solution of the initial trajectory problem, see Theorem 3.5.2.

A first aim of this section is to give easy to check conditions which ensure that all solutions of the switched DAE (4.1.1) are impulse free. In electrical circuits, impulses occur as sparks and often lead to the destruction of some components, therefore it is important to analyze circuits with respect to the ability to produce impulses. Furthermore, switches might be induced by faults, hence the switching signal is not known and therefore the results will be independent of the switching signal. In addition, a simple condition will be given, which ensures that jumps do not occur in the state variables, i.e. a condition that guaranties that all solutions are actually classical solutions. The conditions for impulse and/or jump freeness of the solutions of (4.1.1) are formulated in terms of so called consistency projectors. It is possible to construct these projectors directly in terms of the matrices (E_p, A_p), $p = 1, \ldots, N$; it is not necessary to explicitly calculate some normal form (see Definition 4.2.5 together with Theorem 4.2.4).

A second aim is to study the stability of switched DAEs (4.1.1). When each matrix E_p is invertible, (4.1.1) reduces to a more familiar switched ordinary differential equation (switched ODE), or switched system. The stability theory of switched ODEs has received considerable attention in the last couple of decades, and is now relatively mature. In particular, it is well known that switching among stable subsystems may lead to instability; a switched system is asymptotically stable under arbitrary switching if (and only if) the subsystems share a common Lyapunov function; and stability is preserved under sufficiently slow switching, as can be shown using multiple Lyapunov functions (one for each subsystem). The reader is referred to the book [Lib03] for these and other results on switched systems and for an extensive literature overview.

On the other hand, an investigation of stability questions for switched DAEs by similar methods has not yet appeared in the literature. In Section 4.3, Lyapunov-based sufficient conditions for stability of switched DAEs are established. In the special case of switched ODEs, the results reduce to the known results mentioned above. However, it will be demonstrated by means of examples that the presence of algebraic constraints leads to new types of instability mechanisms.

4.2 Impulse free solutions

As mentioned above the aim of this section is to give conditions for which the switched DAE (4.1.1) has impulse free solutions. Firstly the so called Quasi-Weierstrass form together with a special sequence of subspaces will be introduced. This makes the definition of the so called consistency projectors possible, which play a fundamental role for the formulation of the sought conditions for impulse free solutions.

4.2.1 The Quasi Weierstraß form

It is well known that for a regular matrix pair $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ there exists invertible matrices $S, T \in \mathbb{C}^{n\times n}$ such that $(SET, SAT) = ([\begin{smallmatrix} I & \\ & N \end{smallmatrix}], [\begin{smallmatrix} J & \\ & I \end{smallmatrix}])$, where $J \in \mathbb{C}^{n_1\times n_1}$, $n_1 \in \mathbb{N}$, is in Jordan canonical form and $N \in \mathbb{C}^{n_2\times n_2}$, $n_2 := n-n_2$, is a nilpotent matrix in Jordan canonical form. However, the proofs of this result are in general not constructive and the matrices S and T cannot be given easily in terms of the original matrices E and A.

In most situations it is not necessary that the matrices J and N are in Jordan canonical form, in fact, this assumption often yield complex-valued matrices which might not be desirable if one starts with real valued matrices E and A. In the following, real-valued invertible matrices $S, T \in \mathbb{R}^{n\times n}$ will be constructed such that the matrix pair (E, A) is put into the *Quasi Weierstraß form*

$$\begin{aligned}
&(SET, SAT) = \left(\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \begin{bmatrix} J & 0 \\ 0 & I \end{bmatrix} \right) \\
&J \in \mathbb{R}^{n_1\times n_1}, n_1 \in \mathbb{N}, \\
&N \in \mathbb{R}^{(n-n_1)\times(n-n_1)} \text{ nilpotent.}
\end{aligned} \tag{4.2.1}$$

Note that it is not assumed that J and N are in any special form, hence the Quasi Weierstraß form is *not a normal form* in the strict sense. For obtaining the Quasi Weierstraß form, the following *Wong sequences* [Won74] of linear subspaces play a fundamental role.

$$\begin{aligned}
\mathcal{V}_0 &= \mathbb{R}^n, & \mathcal{V}_{i+1} &= A^{-1}(E\mathcal{V}_i),\ i = 0, 1, \dots, \\
\mathcal{W}_0 &= \{0\}, & \mathcal{W}_{i+1} &= E^{-1}(A\mathcal{W}_i),\ i = 0, 1, \dots.
\end{aligned} \tag{4.2.2}$$

Note that the Wong sequences are nested, i.e. $\mathcal{V}_{i+1} \subseteq \mathcal{V}_i$ and $\mathcal{W}_{i+1} \supseteq \mathcal{W}_i$.

In the following, several lemmas are given which establish some important properties of the subspaces $\mathcal{V}_i$ and $\mathcal{W}_i$. Most of these result are not stated explicitly in [Won74], however, hidden in a proof, the

Quasi-Weierstraß appears in an implicit form. Furthermore, in [OD85] the first Wong-sequence is studied (without being aware of [Won74]) and some of the following results can be found there as well.

Lemma 4.2.1 (Explicit representation of $\mathcal{V}_i$ and $\mathcal{W}_i$)
Consider a regular matrix pair $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, and let $\mathcal{V}_i$ and $\mathcal{W}_i$, $i \in \mathbb{N}$, be given as in (4.2.2). Then

$$\forall i \in \mathbb{N}\ \forall \lambda \in \mathbb{R} \text{ with } \det(A - \lambda E) \neq 0 : \begin{cases} \mathcal{V}_i = \operatorname{im}\left((A-\lambda E)^{-1}E\right)^i, \\ \mathcal{W}_i = \ker\left((A-\lambda E)^{-1}E\right)^i. \end{cases}$$

□

Proof. Let $\lambda \in \mathbb{R}$ be such that $\det(A - \lambda E) \neq 0$ and let $F := (A - \lambda E)^{-1}E$.

Step 1: $\mathcal{V}_i = \operatorname{im} F^i$ for all $i \in \mathbb{N}$ is proved by induction.
For $i = 0$ the assertion holds by definition, hence assume $\mathcal{V}_i = \operatorname{im} F^i$ holds for some $i \in \mathbb{N}$. Invoking the relations $\mathcal{V}_{i+1} \subseteq \mathcal{V}_i$ and $\operatorname{im} F^{i+1} \subseteq \operatorname{im} F^i$ the following equivalences hold for all $x \in \mathbb{R}^n$:

$$\begin{aligned} x \in \mathcal{V}_{i+1} \quad &\Leftrightarrow \quad \exists\, y \in \mathcal{V}_i :\ Ax = Ey \\ &\Leftrightarrow \quad \exists\, y \in \mathcal{V}_i :\ (A - \lambda E)x = E\underbrace{(y - \lambda x)}_{=:\widehat{y} \in \mathcal{V}^i} \\ &\Leftrightarrow \quad \exists\, \widehat{y} \in \mathcal{V}^i = \operatorname{im} F^i :\ (A - \lambda E)x = E\widehat{y} \\ &\Leftrightarrow \quad \exists\, \widehat{y} \in \operatorname{im} F^i :\ x = F\widehat{y} \\ &\Leftrightarrow \quad x \in \operatorname{im} F^{i+1} \end{aligned}$$

Step 2: $\mathcal{W}_i = \ker F^i$ for all $i \in \mathbb{N}$ is proved by induction.
For $i = 0$ the assertion holds by definition, hence assume $\mathcal{W}_i = \ker F^i$. Observe first that the linear operator $(I + \lambda F) : \ker F^i \to \ker F^i$ is invertible with inverse $-\sum_{j=0}^{i-1}(-\lambda)^j F^j$, therefore the following equi-

valences hold for all $x \in \mathbb{R}^n$:

$$\begin{aligned} x \in W_{i+1} \quad &\Leftrightarrow \quad \exists\, y \in W_i : \; Ex = Ay = (A - \lambda E)y + \lambda E y \\ &\Leftrightarrow \quad \exists\, y \in W_i = \ker F^i : \; Fx = (I + \lambda F)y =: \widehat{y} \\ &\Leftrightarrow \quad \exists\, \widehat{y} \in \ker F^i : \; Fx = \widehat{y} \\ &\Leftrightarrow \quad x \in \ker F^{i+1} \end{aligned}$$ qed

Lemma 4.2.2 (Properties of $\mathcal{W}^*$ and $\mathcal{V}^*$)
Let $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, be a regular matrix pair and let $\mathcal{V}_i$, $\mathcal{W}_i$, $i \in \mathbb{N}$, be given by (4.2.2). Let $k^*, l^* \in \mathbb{N}$ be such that

$$\begin{aligned} \mathcal{V}_0 \supset \mathcal{V}_1 \supset \cdots \supset \mathcal{V}_{k^*} = \mathcal{V}_{k^*+1} = \cdots\,, \\ \mathcal{W}_0 \subset \mathcal{W}_1 \subset \cdots \subset \mathcal{W}_{l^*} = \mathcal{W}_{l^*+1} = \cdots\,. \end{aligned}$$

Let $\mathcal{V}^* := \mathcal{V}_{k^*} = \bigcap_i \mathcal{V}_i$ and $\mathcal{W}^* := \mathcal{W}_{l^*} = \bigcup_i \mathcal{W}_i$, then, clearly, $A\mathcal{V}^* \subseteq E\mathcal{V}^*$, $E\mathcal{W}^* \subseteq A\mathcal{W}^*$ and, furthermore,

(i) $k^* = l^*$,

(ii) $\mathcal{V}^* \oplus \mathcal{W}^* = \mathbb{R}^n$,

(iii) $\ker E \cap \mathcal{V}^* = \{0\}$ and $\ker A \cap \mathcal{W}^* = \{0\}$. □

Proof. Since $\mathcal{V}_i$ and $\mathcal{W}_i$ are linear subspaces of a finite dimensional linear space the values k^* and l^* are well defined.

(i) Let $\lambda \in \mathbb{R}$ be such that $\det(A - \lambda E) \neq 0$ and let $F := (A - \lambda E)^{-1}E$, then, by Lemma 4.2.1 and for all $i \in \mathbb{N}$,

$$\dim \mathcal{V}_i + \dim \mathcal{W}_i = \dim \operatorname{im} F^i + \dim \ker F^i = n, \tag{4.2.3}$$

this implies $k^* = l^*$.

(ii) In view of (4.2.3) it suffices to show that $\mathcal{V}^* \cap \mathcal{W}^* = \{0\}$. Let $x \in \mathcal{V}^* \cap \mathcal{W}^*$ and $F = (A - \lambda E)^{-1}E$ for some $\lambda \in \mathbb{R}$ with $\det(A - \lambda E) \neq 0$. Then, by Lemma 4.2.1 and (i), there exists $y \in \mathbb{R}^n$ such that $x = F^{k^*}y$ and $0 = F^{k^*}x = F^{2k^*}y$. Hence $y \in \ker F^{2k^*} = \mathcal{W}_{2k^*} = \mathcal{W}_{k^*} = \ker F^{k^*}$, i.e. $0 = F^{k^*}y = x$.

(iii) From $\ker A \subseteq \mathcal{V}_i$ and $\ker E \subseteq \mathcal{W}_i$ for all $i \in \mathbb{N}$ it follows that $\ker A \subseteq \mathcal{V}^*$ and $\ker E \subseteq \mathcal{W}^*$, hence (ii) implies the claim. $\boxed{\text{qed}}$

Lemma 4.2.3 (Invertibility of $[EV, AW]$)
For a regular matrix pair $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$, let $\mathcal{V}^*, \mathcal{W}^* \subseteq \mathbb{R}^n$ be given as in Lemma 4.2.2 and let $n_1 := \dim \mathcal{V}^*$. Choose $V \in \mathbb{R}^{n\times n_1}$ and $W \in \mathbb{R}^{n\times(n-n_1)}$ such that $\operatorname{im} V = \mathcal{V}^*$ and $\operatorname{im} W = \mathcal{W}^*$. Then $T := [V, W]$ and $S^{-1} := [EV, AW]$ are invertible matrices. $\square$

Proof. Invertibility and existence of T follows Lemma 4.2.2(ii). To show invertibility of S^{-1} it suffices to show that $\ker[EV, AW] = \{0\}$. Therefore, consider any $x \in \mathbb{R}^n_1$ and $y \in \mathbb{R}^{n-n_1}$ with $EVx = 0$ and $AWy = 0$. Invoking Lemma 4.2.2(iii) yields $Vx \in \mathcal{V}^* \cap \ker E = \{0\}$ and $Wy \in \mathcal{W}^* \cap \ker A = \{0\}$. Since V and W have full column rank it follows that $x = 0$ and $y = 0$. $\boxed{\text{qed}}$

Combining all the above results it is now possible to formulate the main result.

Theorem 4.2.4 (Quasi-Weierstraß form)
Consider a regular matrix pair $(E, A) \in \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$ and let $\mathcal{V}^*, \mathcal{W}^* \subseteq \mathbb{R}^n$ be given as in Lemma 4.2.2, i.e. $\mathcal{V}^* := \bigcap_i \mathcal{V}_i$ and $\mathcal{W}^* := \bigcup_i \mathcal{W}_i$, where $\mathcal{V}_i$ and $\mathcal{W}_i$ are given by (4.2.2). Choose $V \in \mathbb{R}^{n\times n_1}$, $n_1 \in \mathbb{N}$, and $W \in \mathbb{R}^{n\times(n-n_1)}$ such that $\operatorname{im} V = \mathcal{V}^*$ and $\operatorname{im} W = \mathcal{W}^*$. Then $T := [V, W]$ and $S := [EV, AW]^{-1}$ put (E, A) into a Quasi Weierstraß form (4.2.1), i.e.

$$(SET, SAT) = \left(\begin{bmatrix} I & \\ & N \end{bmatrix}, \begin{bmatrix} J & \\ & I \end{bmatrix} \right),$$

where $J \in \mathbb{R}^{n_1\times n_1}$ is some matrix, $N \in \mathbb{R}^{(n-n_1)\times(n-n_1)}$ is a nilpotent matrix. $\square$

Proof. Invertibility and existence of $T, S \in \mathbb{R}^{n\times n}$ was shown in Lemma 4.2.2(ii) and Lemma 4.2.3.

Hence it remains to show that (SET, SAT) is in Quasi Weierstraß form, i.e. it must be shown that there exists a nilpotent matrix $N \in$

$\mathbb{R}^{(n-n_1)\times(n-n_1)}$ and a matrix $J \in \mathbb{R}^{n_1\times n_1}$ such that

$$E[V,W] = [EV, AW]\begin{bmatrix} I & \\ & N\end{bmatrix} \quad \text{and} \quad A[V,W] = [EV, AW]\begin{bmatrix} J & \\ & I\end{bmatrix}$$

or, equivalently, $EW = AWN$ and $AV = EVJ$. The existence of J and N with the latter property follows easily from the relations $E\mathcal{W}^* \subseteq A\mathcal{W}^*$ and $A\mathcal{V}^* \subseteq E\mathcal{V}^*$, hence it remains to show that N is nilpotent.

First observe that by Lemma 4.2.2(iii), for all $x, y \in \mathbb{R}^n$ and for all $i \in \mathbb{N}$,

$$Ex = Ay \ \wedge \ x \in \mathcal{W}_{i+1} \ \wedge \ y \in \mathcal{W}^* \quad \Rightarrow \quad y \in \mathcal{W}_i.$$

Let k^* be such that $\mathcal{W}^* = \mathcal{W}_{k^*}$, then for any $x \in \mathbb{R}^{n-n_1}$

$$\begin{array}{rcl} Wx \in \mathcal{W}_{k^*} & \overset{EWx=AWNx}{\Longrightarrow} & WNx \in \mathcal{W}_{k^*-1} \\ & \overset{EWNx=AWN^2x}{\Longrightarrow} & WN^2x \in \mathcal{W}_{k^*-2} \\ & \cdots & \\ & \overset{EWN^{k^*-1}x=AWN^{k^*}x}{\Longrightarrow} & WN^{k^*}x \in \mathcal{W}_0 = \{0\}, \end{array}$$

since W has full column rank it is therefore shown that $N^{k^*}x = 0$ for all $x \in \mathbb{R}^{n-n_1}$, i.e. N is nilpotent. $\boxed{\text{qed}}$

As already mentioned, it is very simple to calculate the matrices V and W in Theorem 4.2.4, in fact, it is even possible to do the calculation symbolically. In the following an implementation in Matlab is given, where the build-in Matlab functions *colspace* and *null* are used:

Listing 1: Matlab function for calculating a basis of the preimage $A^{-1}(\operatorname{im} S)$ for some matrices A and S

```
function V=getPreImage(A,S)
[m1,n1]=size(A); [m2,n2]=size(S);
if m1==m2 | m2==0
    H=null([A,S]);
    V=colspace(H(1:n1,:));
else
    error('Both matrices must have same number of rows');
end;
```

Listing 2: Matlab function for calculating a basis of the space $\mathcal{V}^*$ as in Theorem 4.2.4

```
function V = getVspace(E,A)
[m,n]=size(E);
if (m==n) & size(E)==size(A)
    V=eye(n,n);
    oldsize=n;
    newsize=n;
    finished=0;
    while finished==0;
       EV=colspace(E*V);
       V=getPreImage(A,EV);
       oldsize=newsize;
       newsize=rank(V);
       finished=(newsize==oldsize);
    end;
else
    error('Matrices E and A must be square and of the same size');
end;
```

Listing 3: Matlab function for calculating a basis of the space $\mathcal{W}^*$ as in Theorem 4.2.4

```
function W = getWspace(E,A)
[m,n]=size(E);
if (m==n) & size(E)==size(A)
    W=zeros(n,1);
    oldsize=0;
    newsize=0;
    finished=0;
    while finished==0;
       AW=colspace(A*W);
       W=getPreImage(E,AW);
       oldsize=newsize;
       newsize=rank(W);
       finished=(newsize==oldsize);
    end;
else
    error('Matrices E and A must be square and of the same size');
end;
```

4.2.2 Consistency projectors

Definition 4.2.5 (Consistency projectors)
For a regular matrix pair (E, A), let $T \in \mathbb{R}^{n\times n}$ and $n_1 \in \mathbb{N}$ be given as in Theorem 4.2.4. The *consistency projector* for the pair (E, A) is

$$\Pi_{(E,A)} := T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^{-1},$$

where $I \in \mathbb{R}^{n_1\times n_1}$ is an identity matrix of size $n_1 \times n_1$. □

Note that the consistency projector does not depend on the specific choice of $T = [V, W]$, because for any other choice $\widehat{T} = [\widehat{V}, \widehat{W}]$ with $\operatorname{im} \widehat{V} = \mathcal{V}^*$ and $\operatorname{im} \widehat{W} = \mathcal{W}^*$ there exists invertible matrices $P \in \mathbb{R}^{n_1\times n_1}$ and $Q \in \mathbb{R}^{n_2\times n_2}$ such that $\widehat{V} = VP$ and $\widehat{W} = WQ$, hence

$$\begin{aligned}\widehat{T} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \widehat{T}^{-1} &= [V, W] \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \left([V, W] \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \right)^{-1} \\ &= [V, W] \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} [V, W]^{-1} = \Pi_{(E,A)}\end{aligned}$$

To show how the consistency projectors are related to solutions of switched DAEs (4.1.1) the following lemma on the nature of local solutions is necessary.

Lemma 4.2.6 (Explicit local solution)
Let Assumptions (S1) and (S2) hold and let $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be some ITP solution of (4.1.1) with initial time $t_0 \in \mathbb{R}$. Furthermore, let $s, t \in \mathbb{R}$ with $t_0 \leq s < t$ be such that the switching signal σ is constant on $[s, t)$. Then there exists an invertible matrix $T \in \mathbb{R}^{n\times n}$, a matrix $J \in \mathbb{R}^{n_1\times n_1}$, $n_1 \in \mathbb{N}$, and $v_0 \in \mathbb{R}^{n_1}$ such that

$$x_{(s,t)} = T \begin{pmatrix} \left(\tau \mapsto e^{J(\tau-s)} v_0\right)_{\mathbb{D}} \\ 0 \end{pmatrix}_{(s,t)} \tag{4.2.4}$$

□

Proof. Let $p := \sigma(s)$. Since x is an ITP solution, it follows that $(E_\sigma \dot{x})_{[t_0,\infty)} = (A_\sigma x)_{[t_0,\infty)}$ and, in particular (see also Proposition 2.4.8),

$$E_p \dot{x}_{(s,t)} = A_p x_{(s,t)}.$$

Choose the matrices T, J and N corresponding to the regular matrix pair (E_p, A_p) as in Theorem 4.2.4. Let $\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right) := T^{-1}x$ where $v \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n_1}$ and $w \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n_2}$, then

$$\begin{aligned} \dot{v}_{(s,t)} &= Jv_{(s,t)} \\ N\dot{w}_{(s,t)} &= w_{(s,t)}. \end{aligned}$$

It remains to show that a) $v_{(s,t)} = \big((\tau \mapsto e^{J(\tau-s)}v_0)_{\mathbb{D}}\big)_{(s,t)}$ for some $v_0 \in \mathbb{R}^{n_1}$ and b) $w_{(s,t)} = 0$.

To show a), first observe that by Proposition 2.2.10 the equation $\dot{v}_{(s,t)} = Jv_{(s,t)}$ is equivalent to

$$\forall \varphi \in \mathcal{C}_0^\infty \text{ with support in } (s,t): \quad \dot{v}(\varphi) = Jv(\varphi).$$

Now [Wal94, 6.VII.Satz] yields that any distributional solution v of $\dot{v} = Jv$ on the open interval (s,t) can be represented by a classical solution of the classical ODE $\dot{v} = Jv$, i.e. a) is shown.

To show b), take the derivative of the equation $N\dot{w}_{(s,t)} = w_{(s,t)}$, restrict it to (s,t) and multiply it from the left with N to obtain

$$N^2\ddot{w}_{(s,t)} = N\dot{w}_{(s,t)}.$$

Note that the differentiation produces Dirac impulses at s and t, however these are deleted by the restriction to the open interval (s,t). This process can be repeated and since N is nilpotent it follows that $N^{n_2} = 0$ where $n_2 := n - n_1$, hence

$$0 = N^{n_2}{w^{(n_2)}}_{(s,t)} = N^{n_2-1}{w^{(n_2-1)}}_{(s,t)} = \ldots = N\dot{w}_{(s,t)} = w_{(s,t)}$$

and b) is shown. $\boxed{\text{qed}}$

Remark 4.2.7 (Represenation of local solutions)
Lemma 4.2.6 only states the existence of matrices $T \in \mathbb{R}^{n\times n}$ and $J \in \mathbb{R}^{n_1\times n_1}$ such that all ITP solutions of (4.1.1) fulfill (4.2.4). However, the proof of Lemma 4.2.6 reveals that (4.2.4) holds for any invertible matrix $T \in \mathbb{R}^{n\times n}$ and any matrix $J \in \mathbb{R}^{n_1\times n_1}$, $n_1 \in \mathbb{N}$, for which there exists an invertible matrix $S \in \mathbb{R}^{n\times n}$ and a nilpotent matrix $N \in \mathbb{R}^{(n-n_1)\times(n-n_1)}$ such that $(SE_{\sigma(s)}T, SA_{\sigma(s)}T) = \left(\left[\begin{smallmatrix} I & \\ & N \end{smallmatrix}\right], \left[\begin{smallmatrix} J & \\ & I \end{smallmatrix}\right]\right)$. $\square$

Finally, it is stressed that the ITP solution x in Lemma 4.2.6 is only considered on the *open* interval (s,t), in particular nothing is said about the impulsive part $x[s]$. With the above results on local solutions it is now possible to prove the following main result on the consistency projectors.

Theorem 4.2.8 (Consistency projectors and solutions)
Consider the switched DAE (4.1.1) with Assumptions (S1) and (S2). For each $p \in \{1,\dots,N\}$, let $\Pi_p := \Pi_{(E_p,A_p)}$ be the consistency projectors as in Definition 4.2.5. Then every ITP solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of (4.1.1) with initial time t_0 fulfills

$$\forall t \geq t_0 : \quad x(t+) = \Pi_{\sigma(t)} x(t-) \qquad \square$$

Proof. Let $p = \sigma(t)$ and for the matrix pair (E_p, A_p) choose the matrices S,T,J,N as in Theorem 4.2.4. By Assumption (S1) there exists $\varepsilon > 0$ such that σ is constant on $[t, t+\varepsilon)$, hence Lemma 4.2.6 (together with Remark 4.2.7) yields

$$x(t+) = T \begin{pmatrix} v_0 \\ 0 \end{pmatrix}$$

for some $v_0 \in \mathbb{R}^{n_1}$, $n_1 \in \mathbb{N}$. Let $T^{-1}x(t-) = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$, where $x_1 \in \mathbb{R}^{n_1}$ and $x_2 \in \mathbb{R}^{n-n_1}$. Then

$$\Pi_{\sigma(t)} x(t-) = T \begin{pmatrix} x_1 \\ 0 \end{pmatrix},$$

so it remains to show that $x_1 = v_0$.

Let $T^{-1}x = \left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$, then $v(t-) = x_1$ and $v(t+) = v_0$ and, since x is an ITP solution of (4.1.1),

$$E_p \dot{x}_{[t,t+\varepsilon)} = A_p x_{[t,t+\varepsilon)},$$

multiplying from the left with S and substituting x by $T\left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$ yield

$$\dot{v}_{[t,t+\varepsilon)} = J v_{[t,t+\varepsilon)}.$$

It remains to show that $v(t-) = v(t+)$. Restricting the last differential equation to the point t, i.e. considering the impulsive part of it, gives $\dot{v}[t] = Jv[t]$ and since $v[t]$ is a distribution with point support there exists $a_0, a_1, \ldots, a_N \in \mathbb{R}^{n_1}$, $K \in \mathbb{N}$, such that

$$v[t] = a_0\delta_t + a_1\delta_t' + \ldots + a_K\delta^{(n)},$$

hence, invoking Corollary 2.3.5,

$$\big(v(t+) - v(t-)\big)\delta_t + \sum_{k=0}^{K} a_k\delta_t^{(k+1)} = \sum_{k=0}^{K} a_k\delta_t^{(k)},$$

or

$$0 = \sum_{k=0}^{K+1} b_k\delta_t^{(k)},$$

where $b_{N+1} = a_N$, $b_k = a_{k-1} - a_k$, $k = N, \ldots, 1$, and $b_0 = v(t+) - v(t-) - a_0$. Since $\delta_t, \delta_t', \ldots, \delta_t^{(N+1)}$ are linearly independent it follows that $0 = b_{N+1} = \ldots = b_0$. Hence $0 = a_N = \ldots a_0 = 0$ and finally $v(t+) - v(t-) = 0$ which completes the proof. qed

Combining Lemma 4.2.6 and Remark 4.2.7 with Theorem 4.2.8 immediately gives the following corollary.

Corollary 4.2.9 (No jumps if no switches)
Consider the switched DAE (4.1.1) with assumptions (S1) and (S2) and let $x \in (\mathbb{D}_{\text{pw}\mathcal{C}^\infty})^n$ be an ITP solution of (4.1.1) with initial time $t_0 \in \mathbb{R}$. Then

$$\forall t > t_0: \quad \sigma(t-) = \sigma(t+) \;\Rightarrow\; x(t+) = x(t-),$$

i.e. jumps in the solutions can only occur at switching times or at the initial time t_0. □

4.2.3 Sufficient conditions for impulse/jump freeness of solutions

In general, a solution of (4.1.1) will have jumps and impulses. In the following, sufficient conditions will be given which ensure that every solution of (4.1.1) under arbitrary switching is impulse free or, additionally, has no jumps.

Assumptions
For the switched DAE (4.1.1) and $p = 1, \ldots, N$, let $\Pi_p := \Pi_{(E_p, A_p)}$ be the consistency projectors as in Definition 4.2.5.

(A1) $\forall p \in \{1, \ldots, N\}: \; E_p(I - \Pi_p) = 0$ or

(A2) $\forall p, q \in \{1, \ldots, N\}: E_p(I - \Pi_p)\Pi_q = 0$ or

(A3) $\forall p, q \in \{1, \ldots, N\}: \; (I - \Pi_p)\Pi_q = 0.$ □

Note that the above three assumptions are alternatives, i.e. the following results will only use one of these assumptions. Since the consistency projectors Π_p can easily be calculated by a finite sequence of subspaces (see Theorem 4.2.4 and Definition 4.2.5) only depending on the original matrix pairs (E_p, A_p), the Assumptions (A1)-(A3) can be checked directly in terms of the original data. The following theorems state the properties of the solutions if one of the Assumptions (A1)-(A3) is fulfilled.

Theorem 4.2.10 (A1)
Consider the switched DAE (4.1.1) satisfying Assumptions (S1), (S2) and (A1). Then, for every impulse free initial trajectory and any initial time, the unique ITP solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ is impulse free, i.e. $x[t] = 0$ for all $t \in \mathbb{R}$ or, in other words, the distributional solution is actually a piecewise-smooth function. □

Proof. Let $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be the ITP solution to some given initial trajectory and initial time $t_0 \in \mathbb{R}$ and let $t_0 < t_1 < t_2 < \ldots$ be the switching times of the switching signal σ after the initial time t_0. Lemma 4.2.6 already shows that $x_{(t_i, t_{i+1})}$ is impulse free for all $i \in \mathbb{N}$, hence it remains to show that $x[t_i] = 0$ for all $i \in \mathbb{N}$. Therefore, consider a fixed $i \geq 0$ and let $p = \sigma(t_i)$. For the matrix pair (E_p, A_p), choose matrices S, T, J, N as in Theorem 4.2.4, i.e. $(SE_pT, SA_pT) = \left(\left[\begin{smallmatrix} I & \\ & N \end{smallmatrix}\right], \left[\begin{smallmatrix} J & \\ & I \end{smallmatrix}\right]\right)$ and let $T^{-1}x = \left(\begin{smallmatrix} v \\ w \end{smallmatrix}\right)$. Then $x[t_i] = 0$ if and only if $v[t_i] = 0$ and $w[t_i] = 0$, where v and w fulfill

$$\begin{aligned} \dot{v}[t_i] &= Jv[t_i], \\ N\dot{w}[t_i] &= w[t_i]. \end{aligned}$$

In the proof of Theorem 4.2.8 it was already shown that $\dot{v}[t_i] = Jv[t_i]$ implies $v[t_i] = 0$. Hence it remains to show that $N\dot{w}[t_i] = w[t_i]$ together with Assumption (A1) implies $w[t_i] = 0$. First observe that $N\dot{w}[t_i] = w[t_i]$ implies, invoking Corollary 2.3.5,

$$N(w[t_i])' = w[t_i] - N\big(w(t_i+) - w(t_i-)\big)\delta_{t_i},$$

taking the derivative of the equations and multiplying it from the left with N yields

$$\begin{aligned} N^2(w[t_i])'' &= N(w[t_i])' - N^2\big(w(t_i+) - w(t_i-)\big)\delta'_{t_i} \\ &= w[t_i] - N\big(w(t_i+) - w(t_i-)\big)\delta_{t_i} - N^2\big(w(t_i+) - w(t_i-)\big)\delta'_{t_i}. \end{aligned}$$

Repeating this process yields, since N is nilpotent,

$$0 = N^{n_1}(w[t_i])^{(n_1)} = w[t_i] - \sum_{k=0}^{n_1-1} N^{k+1}\big(w(t_i+) - w(t_i-)\big)\delta_{t_i}^{(k)}$$

or

$$w[t_i] = \sum_{k=0}^{n_1-1} N^{k+1}\big(w(t_i+) - w(t_i-)\big)\delta_{t_i}^{(k)}.$$

Assumption (A1) and Theorem 4.2.8 yield

$$\begin{aligned} 0 \overset{\text{(A1)}}{=} E_p(I - \Pi_p)x(t_i-) &\overset{\text{Thm. 4.2.8}}{=} E_p\big((x(t_i-) - x(t_i+)\big) \\ &= S\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}\begin{pmatrix} v(t_i-) - v(t_i+) \\ w(t_i-) - w(t_i+) \end{pmatrix} \end{aligned}$$

and, in particular,

$$0 = N\big(w(t_i-) - w(t_i+)\big), \tag{4.2.5}$$

hence $w[t_i] = 0$. $\boxed{\text{qed}}$

Remark 4.2.11 ((A1) implies index one)
Lemma 4.2.6 reveals that $w(t_0+) = 0$ in (4.2.5) and since $w(t_0-)$ can be arbitrary (because the initial trajectory can be arbitrary) Assumption (A1) implies that $N = 0$, i.e. the corresponding matrix pair (E, A) must have index one or less [KM06, Def. 2.9]. □

Note that Assumption (A1) is actually stronger related to impulse free solutions of ITPs than to impulse free solutions of switched systems. In fact, with the help of the proof of Theorem 4.2.10 it is easy to see that Assumption (A1) is equivalent to the condition that each ITP for $E_p\dot{x} = Ax$, $p = 1, \ldots, N$, only has impulse free solutions. In this sense Assumption (A1) is also a necessary condition for impulse free ITP solutions.

Remark 4.2.12 (Jump freeness implies impulse freeness)
Consider a solution x of the switched DAE (4.1.1). Assume that $x(t+) = x(t-)$ for some $t \in \mathbb{R}$ then the proof of Theorem 4.2.10 shows $x[t] = 0$. Hence jump freeness of a solution always implies impulse freeness of this solution. □

Theorem 4.2.13 (A2)
Consider the switched DAE (4.1.1) satisfying Assumptions (S1), (S2) and (A2). Then every consistent solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of (4.1.1) is impulse free, i.e. $x[t] = 0$ for all $t \in \mathbb{R}$. □

Proof. Let $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be some consistent solution of (4.1.1) for some switching signal $\sigma \in \mathcal{S}$. Using the same notation as in the proof of Theorem 4.2.10 the proof can be repeated identically up to where Assumption (A1) is used.

Let $q = \sigma(t_i-)$ and choose for the matrix pair (E_q, A_q) the matrices S_q, T_q, J_q, N_q and $n_{1,q} \in \mathbb{N}$ as in Theorem 4.2.4. Then Lemma 4.2.6 applied to the interval $(t_i - \varepsilon, t_i)$ for sufficiently small $\varepsilon > 0$ yields that there exists some $v_q \in \mathbb{R}^{n_{1,q}}$ such that

$$x(t_i-) = T_q \begin{bmatrix} v_q \\ 0 \end{bmatrix} = T_q \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T_q^{-1} T_q \begin{bmatrix} v_q \\ 0 \end{bmatrix} = \Pi_q x(t_i-).$$

Therefore, Assumption (A2) implies

$$0 = E_p(I - \Pi_p)\Pi_q x(t_i-) = E_p(I - \Pi_p)x(t_i-)$$

and the claim follows as in the proof of Theorem 4.2.10. $\boxed{\text{qed}}$

Remark 4.2.14 ((A2) and an ODE subsystem)
In general, Assumption (A2) is independent of the index of the matrix pairs (E_p, A_p), $p = 1, \ldots, N$. However, if the index of one matrix pair (E_q, A_q), $q \in \{1, \ldots, N\}$, is zero, i.e. E_q is invertible and (E_q, A_q) is an ODE, then the consistency projector Π_q is the identity matrix and Assumption (A2) is equivalent to Assumption (A1). ◻

In view of Remark 4.2.11, Assumption (A2) is much more "suited" for switched DAEs because it also uses the additional information that at a switch the trajectory cannot be arbitrary but must be within the consistency space of the previous DAE. An analogon of Theorem 4.2.10 for jump freeness is not very meaningful because every inconsistent initial value produces a jump, so that only invertibility of E_p, $p = 1, \ldots, N$ can assure jump freeness. The above idea which made (A2) more suitable for switched DAEs than (A1) leads to (A3) as a meaningful condition for jump freeness of switched DAEs.

Theorem 4.2.15 (A3)
Consider the switched DAE (4.1.1) satisfying Assumptions (S1), (S2) and (A3). Then every consistent solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of (4.1.1) is impulse free and has no jumps, i.e. $x[t] = 0$ and $x(t-) = x(t+)$ for all $t \in \mathbb{R}$ or in other words, the distribution x is actually an absolutely continuous function. ◻

Proof. Since Assumption (A3) implies Assumption (A2), it already follows from Theorem 4.2.13 that all solutions of (4.1.1) are impulse free, hence it remains to show that all solutions have no jumps, i.e. every solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ fulfills $x(t-) = x(t+)$ for all $t \in \mathbb{R}$. Let $\sigma \in \mathcal{S}$ be the switching signal of (4.1.1), $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ an arbitrary solution of (4.1.1), $t \in \mathbb{R}$, $q := \sigma(t-)$ and $p := \sigma(t+)$. If $p = q$, then Corollary 4.2.9 already shows that $x(t-) = x(t+)$, hence it remains to consider $p \neq q$. Identically as in the proof of Theorem 4.2.13, it follows that $\Pi_q x(t-) = x(t-)$, hence Assumption (A3) together with Theorem 4.2.8 yield

$$0 = (I - \Pi_p)\Pi_q x(t-) = (I - \Pi_p)x(t-) = x(t-) - x(t+).$$ qed

Remark 4.2.16 (Special switching signal)
The proofs of Theorem 4.2.13 and Theorem 4.2.15 reveals that one can also consider a fixed parameter pair $(p, q) \in \{1, \ldots, N\}^2$. If Assumption (A2) or Assumption (A3) holds only for this pair (p, q), then it follows that the switch from system (E_q, A_q) to system (E_p, A_p) cannot produce impulses or jumps, respectively. This observation can be used to prove impulse and jump freeness of solutions of the switched DAE 4.1.1 even if Assumptions (A2) or (A3) are not fulfilled for all pairs $p, q \in \{1, \ldots, N\}$ provided more is known about the switching signal. □

4.2.4 Application to a dual redundant buck converter

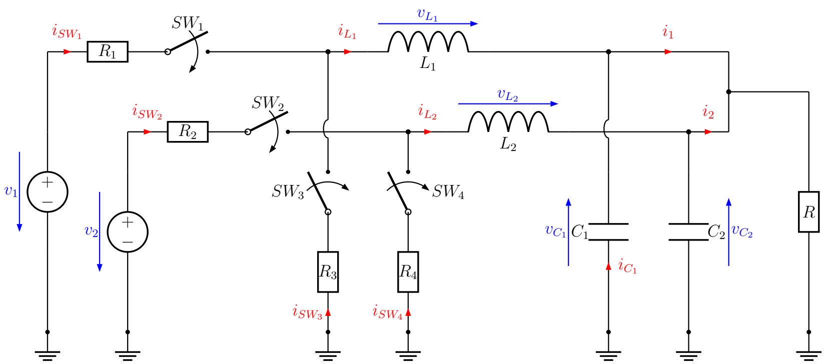

Figure 1: Dual redundant buck converter

Consider the dual-redundant buck converter as illustrated in Figure 1. The purpose of this redundant design is to ensure reliable power delivery to the load R even in the presence of component faults. The two fault free configurations are the "ON" configuration, where the switches SW_1 and SW_2 are closed and the switches SW_3 and SW_4 are open, and the "OFF" configuration where SW_1 and SW_2 are open and SW_3 and SW_4 are closed, in particular, all four switches are synchronized.

If the converter is properly designed, the switching between the nominal configurations should not cause any impulse in the converter state variables. This might not be the case in the presence of component faults. In this regard, a fault in a component will cause a sudden switch from one of the nominal configurations to a faulty configuration. Depending on the nature of the fault, this could induce some of the state variables to suddenly jump or even experience an impulse. This phenomenon could affect some parts of the converter that were not affected by the original fault, potentially destroying the design redundancy, and causing the converter to fail after a single initiating event.

It is now of interest how the circuit behaves in the following fault scenarios: a) some of the switches get stuck in a fixed position, b) a short-circuit occurs in C_1. As common state variables for all configuration choose

$$x=(v_1,v_2,v_{L_1},v_{L_2},v_{C_1},v_{C_2},i_{L_1},i_{L_2},i_{SW_1},i_{SW_2},i_{SW_3},i_{SW_4},i_1,i_2,i_{C_1})^\top.$$

where v_1, v_2 are the input voltages, modelled as constant state variables by $\frac{\mathrm{d}}{\mathrm{d}t}v_1 = 0 = \frac{\mathrm{d}}{\mathrm{d}t}v_2$, the variables v_{L_1}, v_{L_2}, v_{C_1}, v_{C_2} stand for the voltages of the inductors and capacitors, i_{L_1}, i_{L_2}, i_{SW_1}, i_{SW_2}, i_{SW_3}, i_{SW_4}, i_{C_1} are the currents through the switches, inductors and capacitor C_1, finally, i_1, i_2 are the currents which add up to the current through the load R.

The following equations hold independently of the position of the switches:

$$\begin{aligned} L_1\tfrac{\mathrm{d}}{\mathrm{d}t}i_{L_1} &= v_{L_1}, \quad i_{L_1} = i_{SW_1} + i_{SW_3},\\ L_2\tfrac{\mathrm{d}}{\mathrm{d}t}i_{L_2} &= v_{L_2}, \quad i_{L_2} = i_{SW_2} + i_{SW_4} \end{aligned}$$

and

$$\begin{aligned} i_{C_1} &= i_{L_1} - i_1, \quad v_{C_1} = v_{C_2},\\ C_2\tfrac{\mathrm{d}}{\mathrm{d}t}v_{C_2} &= i_{L_2} - i_2, \quad v_{C_2} = R(i_1 + i_2). \end{aligned}$$

If one of the switches is open, then the corresponding current is zero, otherwise the following holds

$$
\begin{aligned}
SW_1 \text{ closed:}\quad & 0 = v_1 + R_{SW_1} i_{SW_1} + v_{L_1} + v_{C_1}, \\
SW_2 \text{ closed:}\quad & 0 = v_2 + R_{SW_2} i_{SW_2} + v_{L_2} + v_{C_2}, \\
SW_3 \text{ closed:}\quad & 0 = R_{SW_3} i_{SW_3} + v_{L_1} + v_{C_1}, \\
SW_4 \text{ closed:}\quad & 0 = R_{SW_4} i_{SW_4} + v_{L_2} + v_{C_2},
\end{aligned}
$$

where it was assumed that the switches are non-ideal, in the sense that they are behaving as resistors when they are closed, otherwise the switches are assumed to be ideal.

Finally, if the capacitor C_1 is not short-circuited, the equation $i_{C_1} = C_1 \frac{\mathrm{d}}{\mathrm{d}t} v_{C_1}$ holds, otherwise, $v_{C_1} = 0$ holds.

These equations directly yield 32 matrix pairs (E_p, A_p), $p = 0, \ldots, 31$. Let $0 = 1, \ldots, 15$ denote the configurations where C_1 is not short-circuited and furthermore identify each switching position with a binary quadruple and the corresponding number, i.e. all switches open correspond to the quadruple $(0,0,0,0)$ and number 0, only switch SW_2 closed correspond to $(0,1,0,0)$ and the number 4 and all switches closed correspond to $(1,1,1,1)$ and the number 15. For the configurations where C_1 is short-circuited just add 16 to these numbers. The nominal "ON" configuration is then given by

$$
(E_{12}, A_{12}) = \left(
\begin{bmatrix}
1 & & & & & & & & & & & & & & \\
 & 1 & & & & & & & & & & & & & \\
 & & & & & & L_1 & & & & & & & & \\
 & & & & & & & L_2 & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & C_2 & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & C_1 & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & & \\
 & & & & & & & & & & & & & &
\end{bmatrix},
\begin{bmatrix}
0 & & & & & & & & & & & & & & \\
 & 0 & & & & & & & & & & & & & \\
 & & 1 & & & & & & & & & & & & \\
 & & & 1 & & & & & & & & & & & \\
 & & & & & & 1 & & & & & & \text{-}1 & & \text{-}1 \\
 & & & & & & & 1 & & & & & & \text{-}1 & \\
 & & & & & & 1 & & \text{-}1 & & \text{-}1 & & & & \\
 & & & & & & & 1 & & \text{-}1 & & \text{-}1 & & & \\
 & & & & 1 & \text{-}1 & & & & & & & & & \\
 & & & & & 1 & & & & & & & \text{-}R & \text{-}R & \\
 & & & & & & & & & & & & & & 1 \\
1 & & 1 & & 1 & & & & R_1 & & & & & & \\
 & 1 & & 1 & & 1 & & & & R_2 & & & & & \\
 & & & & & & & & & & 1 & & & & \\
 & & & & & & & & & & & 1 & & &
\end{bmatrix}
\right)
$$

and the "OFF" configuration is

$$
(E_3, A_3) = \left(\begin{bmatrix}
1 &&&&&&&&&&&&&& \\
& 1 &&&&&&&&&&&&& \\
&&&&&&& L_1 &&&&&&& \\
&&&&&&&& L_2 &&&&&& \\
&&&&&& C_2 &&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&& C_1 &&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&& \\
&&&&&&&&&&&&&&
\end{bmatrix}, \begin{bmatrix}
0 &&&&&&&&&&&&&& \\
& 0 &&&&&&&&&&&&& \\
&& 1 &&&&&&&&&&&& \\
&&& 1 &&&&&&&&&&& \\
&&&&&& 1 &&&&&& \text{-}1 && \text{-}1 \\
&&&&&&& 1 &&&&&& \text{-}1 & \\
&&&&&& 1 && \text{-}1 && \text{-}1 &&&& \\
&&&&&&& 1 && \text{-}1 && \text{-}1 &&& \\
&&&& 1 & \text{-}1 &&&&&&&&& \\
&&&&& 1 &&&&&&& \text{-}R & \text{-}R & \\
&&&&&&&&&&&&&& 1 \\
&&&&&&&& 1 &&&&&& \\
&&&&&&&&& 1 &&&&& \\
&& 1 && 1 &&&&&& R_3 &&&& \\
&&& 1 && 1 &&&&&& R_4 &&&
\end{bmatrix} \right)
$$

It is easy to check (e.g. with a short Matlab program) that all matrix pairs (E_0, A_0), (E_1, A_1), ..., (E_{31}, A_{31}) are regular by calculating the polynomials $\det(sE_p - A)$, $p = 0, \dots, 31$. With the proposed method in Section 4.2.1, it is not difficult to calculate the consistency projectors as in Definition 4.2.5 and to check the Assumptions (A1), (A2) or (A3). As an illustration, the calculations are done for the nominal "ON" configuration (E_{12}, A_{12}), where $[V_{12}, W_{12}] =$

$$
\left[\begin{array}{ccccc|cccccccccc}
R_1 &&&&&&&&&&&&&& \\
& R_2 &&&&&&&&&&&&& \\
&& R_1 &&& 1 &&&&&&&&& \\
&&& R_2 &&& 1 &&&&&&&& \\
&&&& R_1R_2 &&& C_2 &&&&&&& \\
&&&& R_1R_2 &&& \text{-}C_1 &&&&&&& \\
-1 && -1 && -R_2 &&&&&&&&&& \\
& -1 && -1 & -R_1 &&&&&&&&&& \\
-1 && -1 && -R_2 &&&& 1 &&&&&& \\
& -1 && -1 & -R_1 &&&&& 1 &&&&& \\
&&&&&&&&&& 1 &&&& \\
&&&&&&&&&&& 1 &&& \\
\frac{-C_2}{C_1+C_2} & \frac{C_1}{C_1+C_2} & \frac{-C_2}{C_1+C_2} & \frac{C_1}{C_1+C_2} & \frac{R_1C_1(R+R_2)-R_2C_2R}{R(C_1+C_2)} &&&&&&&& 1 && \\
\frac{C_2}{C_1+C_2} & \frac{-C_1}{C_1+C_2} & \frac{C_2}{C_1+C_2} & \frac{-C_1}{C_1+C_2} & \frac{R_2C_2(R+R_1)-R_1C_1R}{R(C_1+C_2)} &&&&&&&&& 1 & \\
\frac{-C_1}{C_1+C_2} & \frac{-C_1}{C_1+C_2} & \frac{-C_1}{C_1+C_2} & \frac{-C_1}{C_1+C_2} & \frac{C_1(RR_1+RR_2+R_1R_2)}{-R(C_1+C_2)} &&&&&&&&&& 1
\end{array} \right]
$$

and the corresponding consistency projector is

$$
\Pi_{12} = [V_{12}, W_{12}] \left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix} \right] [V_{12}, W_{12}]^{-1}
$$

$$
= \left[\begin{array}{ccccccccccccccc}
1 \\
& 1 \\
-1 &&&& \frac{-C_1}{C_1+C_2} & \frac{-C_2}{C_1+C_2} & -R_1 \\
& -1 &&& \frac{-C_1}{C_1+C_2} & \frac{-C_2}{C_1+C_2} && -R_2 \\
&&&& \frac{C_1}{C_1+C_2} & \frac{C_2}{C_1+C_2} \\
&&&& \frac{C_1}{C_1+C_2} & \frac{C_2}{C_1+C_2} \\
&&&&&& 1 \\
&&&&&&& 1 \\
&&&&&& 1 \\
&&&&&&& 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
&&&& \frac{C_1^2}{R\,(C_1+C_2)^2} & \frac{C_1\,C_2}{R\,(C_1+C_2)^2} & \frac{C_2}{C_1+C_2} & \frac{-C_1}{C_1+C_2} \\
&&&& \frac{C_1\,C_2}{R\,(C_1+C_2)^2} & \frac{C_2^2}{R\,(C_1+C_2)^2} & \frac{-C_2}{C_1+C_2} & \frac{C_1}{C_1+C_2} \\
&&&& \frac{-C_1^2}{R\,(C_1+C_2)^2} & \frac{-C_1\,C_2}{R\,(C_1+C_2)^2} & \frac{C_1}{C_1+C_2} & \frac{C_1}{C_1+C_2}
\end{array}\right].
$$

It now follows that $E_{12}(I - \Pi_{12})$ is *not* the zero matrix hence Assumption (A1) is not fulfilled for the nominal "ON" configuration. In fact, there exists initial values such that the corresponding ITP for (E_{12}, A_{12}) has impulsive solutions. However, in nominal conditions, the initial trajectory for the "ON" configuration is not arbitrary but is given by the "OFF" configuration. The consistence projector Π_3 for the "OFF" configuration (E_3, A_3) is given by

$$
\Pi_3 = \left[\begin{array}{ccccccccccccccc}
1 \\
& 1 \\
&&&& \frac{-C_1}{C_1+C_2} & \frac{-C_2}{C_1+C_2} & -R_3 \\
&&&& \frac{-C_1}{C_1+C_2} & \frac{-C_2}{C_1+C_2} && -R_4 \\
&&&& \frac{C_1}{C_1+C_2} & \frac{C_2}{C_1+C_2} \\
&&&& \frac{C_1}{C_1+C_2} & \frac{C_2}{C_1+C_2} \\
&&&&&& 1 \\
&&&&&&& 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
&&&&&& 1 \\
&&&&&&& 1 \\
&&&& \frac{C_1^2}{R(C_1+C_2)^2} & \frac{C_1C_2}{R(C_1+C_2)^2} & \frac{C_2}{C_1+C_2} & \frac{-C_1}{C_1+C_2} \\
&&&& \frac{C_1C_2}{R(C_1+C_2)^2} & \frac{C_2^2}{R(C_1+C_2)^2} & \frac{-C_2}{C_1+C_2} & \frac{C_1}{C_1+C_2} \\
&&&& \frac{-C_1^2}{R(C_1+C_2)^2} & \frac{-C_1C_2}{R(C_1+C_2)^2} & \frac{C_1}{C_1+C_2} & \frac{C_1}{C_1+C_2}
\end{array}\right].
$$

It turns out that Assumption (A2) is fulfilled for the two nominal configurations $\{12, 3\}$, hence arbitrary switching between the "ON" and "OFF" configurations does not yield impulses in any solutions. However, Assumption (A3) is not fulfilled, i.e. jumps in certain state variables cannot be excluded, in fact, it is easy to see, that a switch to the "OFF" configuration forces the currents i_{SW_1} and i_{SW_2} to zero immediately.

To check whether and which faulty configurations can induce impulses in the state variables the condition (A2) must be check for each pair $p, q \in \{0, \ldots, 31\}$, the result of this check is given in the matrix $\mathcal{I} \in \{\square, \blacksquare\}^{32\times 32}$ given in Figure 2, where $\mathcal{I}_{q,p} = \square$, $p, q \in \{0, \ldots, 31\}$, if, and only if, no impulse can occur at a switch from configuration q to configuration p.

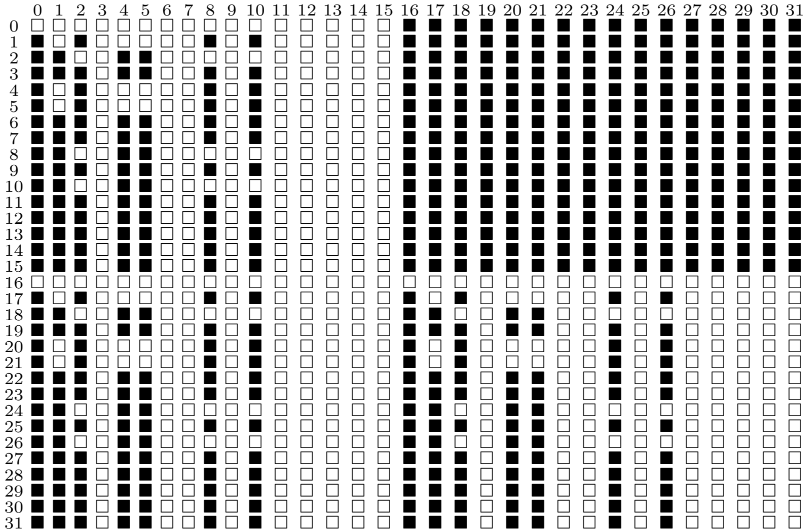

Figure 2: Impulse matrix $\mathcal{I}$ for the configurations $(E_0, A_0), \ldots, (E_{31}, A_{31})$, here $\mathcal{I}_{q,p} = \square$ if, and only if, a switch from configuration q to configuration p cannot produce impulses.

Of special interest is row 12 in $\mathcal{I}$ corresponding to the nominal "ON" configuration, because it can be seen clearly, which faulty configurations might produce impulses in the solution. For example, when a faulty switch back to the "OFF" configuration occurs in the sense that the switches SW_1 and SW_2 are opened first and switches SW_3 and

SW_4 are closed with a small delay (i.e. going from configuration 12 to configuration 3 via the faulty configuration 0), then impulses can occur. On the other hand if in the same situation switches SW_3 and SW_4 are closed first and switches SW_1 and SW_2 are opened later (i.e. going from 12 to 3 via 15) no impulses can occur. Furthermore, the matrix $\mathcal{I}$ reveals that a short cut of the capacitor C_1 (i.e. a switch from some configuration $0, \ldots, 15$ to some configuration $16, \ldots, 31$) can always produce impulses.

4.3 Stability of switched DAEs

In this section the stability of the switched DAE (4.1.1) will be studied. As a first step, it is highlighted that for classical DAEs $E\dot{x} = Ax$ asymptotic stability is equivalent to the existence of a Lyapunov function. For a switched DAE (4.1.1) where each DAE $E_p\dot{x} = A_p x$, $p \in \{1, \ldots, N\}$ is asymptotically stable, sufficient conditions in terms of Lyapunov functions are given which ensure that the switched DAE remains stable under arbitrary switching or under switching with sufficiently large enough dwell time.

Different to classical ODEs, the so called consistency space and its corresponding consistency projectors play a fundamental role in the stability properties of the switched DAE. To illustrate the different nature of switched DAEs several examples are given after the basic definitions and before the formulation of the main results.

4.3.1 Lyapunov functions for classical differential algebraic equations

Consider the classical DAE

$$E\dot{x} = Ax, \tag{4.3.1}$$

where the matrix pair $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ is regular, i.e. $\det(Es - A)$ is not the zero polynomial. A (classical) solution of (4.3.1) is any differentiable function $x : \mathbb{R} \to \mathbb{R}^n$ such that (4.3.1) is fulfilled.

Definition 4.3.1 (Consistency space)
Let the *consistency space* of (4.3.1) be given by

$$\mathfrak{C}_{(E,A)} := \left\{ x^0 \in \mathbb{R}^n \;\middle|\; \begin{array}{l} \exists \text{ solution } x \text{ of } (4.3.1) \\ \text{with } x(0) = x^0 \end{array} \right\}.$$

□

It is well known that for regular matrix pairs each solution of (4.3.1) is uniquely determined by any consistent initial condition $x(0) = x^0 \in \mathfrak{C}_{(E,A)}$. Since (4.3.1) is time invariant, all solutions x evolve within the consistency space, i.e. $x(t) \in \mathfrak{C}_{(E,A)}$ for all $t \in \mathbb{R}$. Furthermore, if (4.3.1) is an ordinary differential equation, i.e. $E \in \mathbb{R}^{n\times n}$ is an invertible matrix, then $\mathfrak{C}_{(E,A)} = \mathbb{R}^n$.

The following lemma gives a nice characterization of the consistency space in terms of the matrices E, A.

Lemma 4.3.2 ([OD85])
Consider the DAE (4.3.1) with regular matrix pair (E, A) and let $\mathcal{V}^*$ be given as in Lemma 4.2.2. Then $\mathfrak{C}_{(E,A)} = \mathcal{V}^*$. In particular, $\ker E \cap \mathfrak{C}_{(E,A)} = \{0\}$. □

A direct consequence of this result is that for the consistency projector $\Pi_{(E,A)}$ as in Definition 4.2.5 the following relation holds:

$$\operatorname{im} \Pi_{(E,A)} = \mathfrak{C}_{(E,A)}. \tag{4.3.2}$$

Definition 4.3.3 (Lyapunov function)
Consider the DAE (4.3.1) with regular matrix pair (E, A) and corresponding consistency space $\mathfrak{C}_{(E,A)} \subseteq \mathbb{R}^n$. Assume there exist a positive definite matrix $P = \overline{P}^\top \in \mathbb{C}^{n\times n}$ and a matrix $Q = \overline{Q}^\top \in \mathbb{C}^{n\times n}$ which is positive definite on $\mathfrak{C}_{(E,A)}$ such that the *generalized Lyapunov equation*

$$A^\top PE + E^\top PA = -Q$$

is fulfilled. Then

$$V : \mathbb{R}^n \to \mathbb{R}_{\geq 0} : x \mapsto (Ex)^\top PEx$$

is called a *Lyapunov function* for the DAE (4.3.1). □

Note that this definition ensures that V is not increasing along solutions, i.e., for any solution $x:\mathbb{R}\to\mathbb{R}^n$ and all $t\in\mathbb{R}$,

$$\tfrac{\mathrm{d}}{\mathrm{d}t}V\big(x(t)\big) = -x(t)^\top Qx(t) \leq 0$$

and equality only holds for $x(t)=0$. Furthermore, the property $\ker E\cap \mathfrak{C}_{(E,A)}=\{0\}$ ensures that V is positive definite on $\mathfrak{C}_{(E,A)}$.

With some abuse of terminology, the DAE (4.3.1) is called *asymptotically stable* if, and only if, $x(t)\to 0$ as $t\to\infty$ for all solutions x of (4.3.1). Note that attractivity of the zero solution already implies attractivity and stability in the sense of Lyapunov for all solutions of (4.3.1), [Ber08]. The following theorem shows the equivalence between asymptotic stability of (4.3.1) and the existence of a Lyapunov function.

Theorem 4.3.4 ([OD85, Ber08])
The DAE (4.3.1) with regular matrix pair (E,A) is asymptotically stable if, and only if, there exists a Lyapunov function $V:\mathbb{R}^n\to\mathbb{R}_{\geq 0}$ for (4.3.1). ◻

Remark 4.3.5
The above definition of a Lyapunov function might seem unsatisfactory because it is not clear how the definition can be generalized to switched DAEs (4.1.1) or non-linear differential algebraic equations. One can say that Definition 4.3.3 is just a "sufficient" definition. Furthermore, a "common Lyapunov function" will be constructed in the proof of Theorem 4.3.9, but it will not precisely be defined what a Lyapunov function for (4.1.1) is. It should be possible to formulate a more general definition of a Lyapunov function for (switched) DAEs and similar results as formulated in the next section will hold, but it would get more technical without adding more insight. ◻

Remark 4.3.6
Assume the invertible matrices $S,T\in\mathbb{R}^{n\times n}$ put the regular matrix pair (E,A) into a Quasi Weierstraß form (4.2.1). It is then easy to see that the classical DAE (4.3.1) is asymptotically stable if, and only if, the *underlying ODE* $\dot v = Jv$ is asymptotically stable, in fact, all solutions of (4.3.1) are given by $x(t)=T\left(\begin{smallmatrix} e^{Jt}v^0\\ 0\end{smallmatrix}\right)$, $t\in\mathbb{R}$, $v^0\in\mathbb{R}^{n_1}$,

$n_1 \in \mathbb{N}$. Hence knowledge of the underlying ODE $\dot{x} = Jx$ and the first n_1 columns of T are sufficient to know everything about the solutions of the classical DAE (4.3.1) and, in particular, its stability properties. $\square$

4.3.2 Switched DAEs: motivating examples

For switched ODEs there exist several well known examples of destabilizing switching. Of course, these are also destabilizing examples for switched DAEs (because every ODE is a special DAE), but in the following, examples are given which are specific to switched DAEs. For the examples, a switching signal $\sigma : \mathbb{R} \to \{1, 2\}$ with a constant interval $\Delta t > 0$ between switching times as illustrated in Figure 3 is considered.

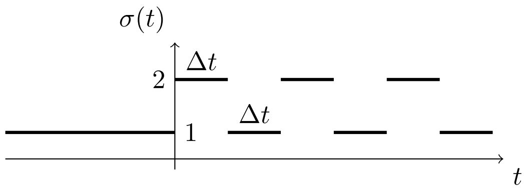

Figure 3: Switching signal $\sigma : \mathbb{R} \to \{1, 2\}$ with constant interval $\Delta t > 0$ between switches.

Example 1a

Let

$$(E_1, A_1) = \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \right),$$
$$(E_2, A_2) = \left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \right).$$

The solutions of the corresponding switched DAE (4.1.1) are shown in Figure 4.

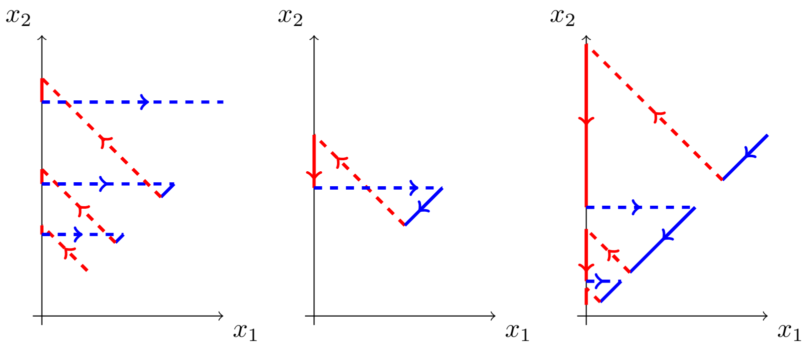

Figure 4: Solutions for Example 1 for different switching signals (dashed lines mean jumps induced by the switching), left: $\Delta t < \frac{1}{2}\ln 2$, all non-trivial solutions grow unbounded; middle: $\Delta t = \frac{1}{2}\ln 2$, all solutions are periodic on $[0,\infty)$; right: $\Delta t > \frac{1}{2}\ln 2$, all solution tend to zero.

For small enough Δt all solutions grow unboundedly and for large enough Δt the solutions converge to zero. Furthermore, there exists a value of Δt for which all solutions are periodic.

The consistency spaces $\mathfrak{C}_p := \mathfrak{C}_{(E_p,A_p)}$, $p = 1,2$ are given by

$$\mathfrak{C}_1 = \operatorname{im}\begin{bmatrix}1\\1\end{bmatrix}, \quad \mathfrak{C}_2 = \operatorname{im}\begin{bmatrix}0\\1\end{bmatrix}.$$

Furthermore, the matrices $T_1 := [V_1, W_1]$ and $T_2 := [V_2, W_2]$ as in Theorem 4.2.4 are

$$T_1 = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}, \quad T_2 = \begin{bmatrix}0 & 1\\1 & -1\end{bmatrix}$$

and the corresponding consistency projectors are

$$\Pi_1 = \begin{bmatrix}0 & 1\\0 & 1\end{bmatrix}, \quad \Pi_2 = \begin{bmatrix}0 & 0\\1 & 1\end{bmatrix}.$$

Note that

$$(E_1T_1, A_1T_2) = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \right) = (E_2T_2, A_2T_2),$$

hence both DAEs are governed by the same underlying scalar ODE $\dot{y} = -y$; in particular, both DAEs are asymptotically stable.

Furthermore, it is easy to see that

$$V : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}, \quad x \mapsto x^\top x$$

restricted to the corresponding consistency space is a Lyapunov function for both subsystems. In spite of this, the switched system is not stable under arbitrary switching.

Example 1b

Let

$$\begin{aligned} (E_1, A_1) &= \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \right), \\ (E_2, A_2) &= \left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right). \end{aligned}$$

The matrix pair (E_1, A_1) is the same as in Example 1a and for the matrix pair (E_2, A_2) the consistency space is the same as in Example 1a, furthermore the matrix $V_2 = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix} \right]$ is the same as in Example 1a.

In view of Remark 4.3.6 it therefore follows that the DAEs $E_p\dot{x} = A_px$, $p = 1, 2$, from Example 1a and 1b have exactly the same solutions.

Nevertheless, the corresponding switched DAEs (4.1.1) are not identical, in fact, they even have opposite stability properties: It was already shown that for Example 1a any switching signal σ as in Figure 3 with sufficiently small Δt destabilizes the switched DAE (4.1.1), on the other hand it is easy to see that the switched DAE 4.1.1 for Example 1b remains asymptotically stable for arbitrary switching, see Figure 5 where the different consistency projectors are illustrated. The consistency projector for (E_2, A_2) is given by

$$\Pi_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

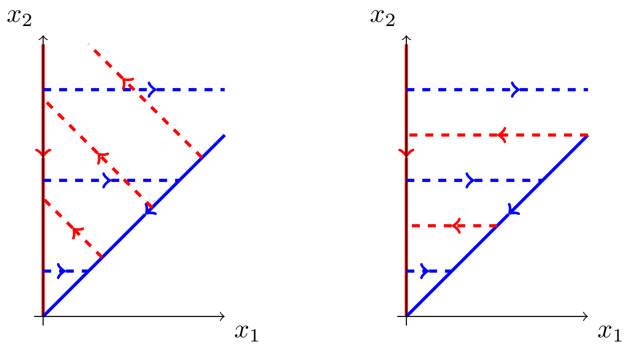

Figure 5: Different jumping behaviour (dashed lines) for Examples 1a and 1b. Left: Example 1a, Right: Example 1b, clearly, the switched DAE (4.1.1) remains asymptotically stable.

Example 2

Let

$$(E_1, A_1) = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 8\pi & 0 \\ \frac{1}{2}\pi & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \right),$$
$$(E_2, A_2) = \left(\begin{bmatrix} 0 & 4 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -4\pi & -4 & 0 \\ -1 & 4\pi & 0 \\ -1 & -4 & 4 \end{bmatrix} \right).$$

The consistency spaces are

$$\mathfrak{C}_1 = \mathbb{R}^3, \quad \mathfrak{C}_2 = \operatorname{im} \begin{bmatrix} 0 & 4 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}$$

and the matrices T_1 and T_2 (as in Definition 4.2.5) are

$$T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 0 & 4 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

With $S_2 = \frac{1}{4}I$ the underlying two dimensional ODE of subsystem 2 is given by

$$\dot{y} = \begin{bmatrix} -1 & -4\pi \\ \pi & -1 \end{bmatrix} y.$$

Selecting $\Delta t = 1/4$ together with a suitable initial condition ensures that the switching only occurs at that moment when the solution is located in the intersection of the consistency spaces (i.e. in $\mathfrak{C}_2$). Hence the solution of the switched DAE exhibits no jumps. The asymptotically stable solutions of the unswitched DAEs are shown in the left part of Figure 6 and the unstable solutions of the switched DAE are illustrated in the right part of Figure 6.

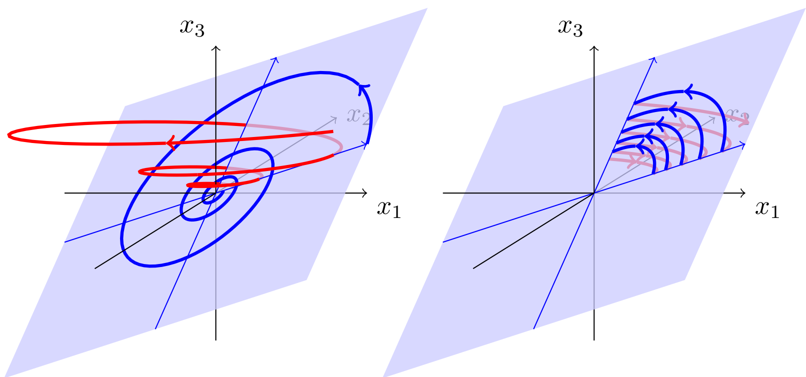

Figure 6: Solutions for the Example 2. Left: Without switching, red: solution of subsystem 1, a three dimensional spiral converging to zero, blue: solution of subsystem 2, a two dimensional spiral converging to zero, Right: with switching, the solutions grow unbounded and exhibit no jumps.

This example illustrates that even in the absence of jumps it is not enough to just study the intersection of the consistency spaces, the unstable behaviour of the switched DAE is basically induced by the solution behaviour of the first subsystem *outside* the intersection of the consistency spaces. However, there does not exists Lyapunov functions

V_1 and V_2 for the two subsystems such that V_1 and V_2 coincide on the intersection of the consistency spaces, because this would imply that all jump free solutions converge to zero (see Corollary 4.3.11).

Example 3

Let

$$(E_1, A_1) = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 2\pi & 0 \\ -2\pi & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right),$$
$$(E_2, A_2) = \left(\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 4\pi & -1 & 4\pi \\ -1 & \pi & -1 \\ 1 & 0 & 0 \end{bmatrix} \right).$$

The consistency spaces are

$$\mathfrak{C}_1 = \operatorname{im} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \mathfrak{C}_2 = \operatorname{im} \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and the matrices T_1 and T_2 (as in Definition 4.2.5) are

$$T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}.$$

The corresponding consistency projectors are then given by

$$\Pi_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

The underlying ODE for the first DAE can be read off directly from the matrix pair (E_1, A_1), the underlying ODE for the second DAE is given by

$$\dot{y} = \begin{bmatrix} -1 & 4\pi \\ -\pi & -1 \end{bmatrix} y.$$

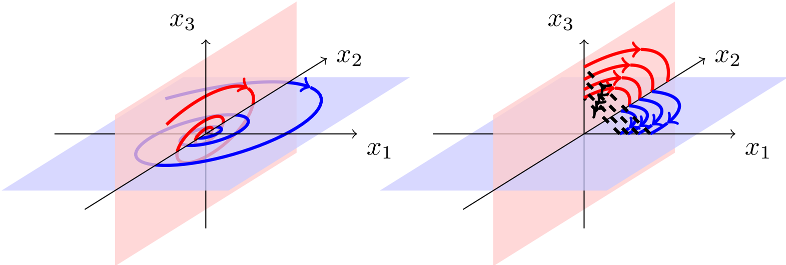

Figure 7: Solutions for the Example 3. Left: Solutions of the individual subsystems, Right: Solutions of the switched system (dashed lines are jumps induced by the switches), the solutions grow unbounded.

The solutions of the unswitched subsystems are illustrated in the left part of Figure 7.

For $\Delta t = 1/2$ and an initial value at $t = 0$ which is located on the x_2 axis, the switching does not induce jumps and all solutions converge to zero. However, the choice $\Delta t = 1/4$ induces jumps and destabilizes the system, see the right part of Figure 7.

Note that $V(x) = x^\top x$ is a common Lyapunov function on the intersection of the consistency spaces. Hence, this example shows that the existence of a common Lyapunov function on the intersection of the consistency space is not sufficient for stability of the switched DAE (4.1.1) under arbitrary switching.

4.3.3 Sufficient conditions for stability of switched DAEs

Definition 4.3.7 (Asymptotic stability)
The switched DAE (4.1.1) is called *asymptotically stable* if, and only if, all distributional solutions are impulse free and each solution $x_{\mathbb{D}}$ given by $x : \mathbb{R} \to \mathbb{R}^n$ fulfills $x(t) \to 0$ as $t \to \infty$. □

What is the relationship of this definition to the classical definition of asymptotic stability (attractivity and stability in the sense of Lyapunov)?

First observe that the impulse freeness assumption is motivated by the stability property, because impulses can be interpreted as peaks with an unbounded height. If there exists a solution with an impulse, then one can scale this solution such that $x(0-)$ gets arbitrarily small, nevertheless, the impulse will remain in the solution, i.e. the solution cannot be interpreted as stable in the classical sense.

Furthermore, linearity implies that it suffices to study the stability properties of the zero solution and the following proposition shows that attractivity of the zero solution implies stability of the zero solution, hence Definition 4.3.7 is justified.

Proposition 4.3.8 (Attractivity implies stability)
Consider the switched DAE (4.1.1) satisfying Assumptions (S1) and (S2). Assume that all solutions of (4.1.1) are impulse free and tend to zero for $t \to \infty$. Then for all $\varepsilon > 0$, $t_0 \in \mathbb{R}$ and all switching signals σ there exists $\delta = \delta(\varepsilon, t_0, \sigma) > 0$ such that the following implication holds

$$x \text{ solves (4.1.1)} \ \wedge \ \|x(0-)\| < \delta \quad \Rightarrow \quad \forall t \geq t_0 : \ \|x(t+)\| \leq \varepsilon,$$

where $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is some norm. □

Proof. For $t_0 \in \mathbb{R}$ let

$$\mathfrak{C}_{t_0} := \left\{ \ x^0 \in \mathbb{R}^n \ \middle| \ \exists \text{ solution } x \text{ of (4.1.1) with } x(t_0-) = x^0 \ \right\},$$

then it is obvious that $\mathfrak{C}_{t_0}$ is a linear subspace of $\mathbb{R}^n$ with some dimension $d \leq n$. If $d = 0$, then nothing is to show, because in this case Theorem 3.5.2 implies that any solution of(4.1.1) is identical to the trivial solution on the interval $[t_0, \infty)$.

Hence assume $d > 0$ and choose a basis $b_1, b_2, \ldots, b_d \in \mathbb{R}^n$ of $\mathfrak{C}_{t_0}$. From Theorem 3.5.2 it follows that there exist unique solutions x^1, x^2, …, $x^d \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of the switched DAE (4.1.1) such that $x^i(t_0-) = b^i$, $i = 1, \ldots, d$, and, by assumption, they all converge to zero for $t \to \infty$.

Furthermore, every solution x of (4.1.1) can be written as a unique linear combination of $x^1, \ldots, x^d$. By assumption each solution is impulse free, hence x^i can be represented by a piecewise-smooth function and, by convergence to zero, x^i is bounded on $[t_0, \infty)$ for all $i = 1, \ldots, d$.

Hence there exists $M_i > 0$ such that $\|x^i(t+)\| \leq M_i$ for all $t \geq t_0$ and $i = 1, \ldots, d$.

For $b \in \mathfrak{C}_{t_0}$ there exist unique $\alpha_1, \ldots, \alpha_d \in \mathbb{R}$ such that $b = \sum_i \alpha_i b_i$, furthermore, the corresponding mapping $b \mapsto \alpha_i =: \alpha_i(b)$ is linear and not identically zero, hence there exists $\overline{\alpha}_i > 0$ such that

$$\left|\alpha_i(b)\right| \leq \overline{\alpha}_i \|b\| \quad \text{for all } i \in \{1, \ldots, d\}.$$

Now choose

$$\delta = \frac{\varepsilon}{\sum_{i=1}^d \overline{\alpha}_i M_i},$$

then for every solution x of (4.1.1) with $\|x(t_0 - \| < \delta$ it follows for all $t \geq t_0$

$$\|x(t+)\| = \left\| \sum_{i=1}^d \alpha_i\big(x(t_0-)\big) x^i(t+) \right\| \leq \sum_{i=1}^d \overline{\alpha}_i \delta M_i = \varepsilon.$$

qed

In addition to the standard assumptions (S1) and (S2) from Section 4.1, throughout this section the Assumption (A2) from Section 4.2.3 is assumed. This ensures that all (consistent) solutions of the switched DAE (4.1.1) are uniquely determined by their past and have no impulses, therefore, all distributional solutions correspond to piecewise-smooth functions.

To simplify the notation in this section, all distributional solutions will be identified with the corresponding piecewise-smooth function.

But it is important to keep in mind that only the distributional framework as introduced in the previous sections allows to speak of *solutions* of the switched DAE (4.1.1), in fact, even in the absence of impulses in the solution x, there may be impulses in $\dot{x}$, hence (4.1.1) makes no sense without the distributional framework.

Theorem 4.3.9 (Stability under arbitrary switching)
Consider the switched DAE (4.1.1) satisfying Assumptions (S1), (S2) and (A2). Let $\Pi_p := \Pi_{(E_p, A_p)} \in \mathbb{R}^{n \times n}$ and $\mathfrak{C}_p := \operatorname{im} \Pi_p$ be the consistency projectors and spaces corresponding to the matrix pairs (E_p, A_p) as in Definition 4.2.5. Assume the classical DAE $E_p \dot{x} = A_p x$ is, for

every $p = 1,\dots,N$, asymptotically stable with Lyapunov function $V_p : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$. If

$$\forall\, p,q \in \{1,\dots,N\}\ \forall x \in \mathfrak{C}_q : \quad V_p(\Pi_p x) \leq V_q(x), \tag{4.3.3}$$

then the switched DAE (4.1.1) is asymptotically stable for every switching signal. ◻

Proof. Theorem 4.2.13 already shows that all (distributional) solutions of (4.1.1) are impulse free, hence it remains to show the convergence to zero.

Step 1: Definition of a common Lyapunov function candidate.
If $x \in \mathfrak{C}_p \cap \mathfrak{C}_q$ for some $p,q \in \{1,\dots,N\}$, then $x = \Pi_p x = \Pi_q x$ hence (4.3.3) implies $V_p(x) = V_q(x)$, therefore

$$V : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \begin{cases} V_p(x), & x \in \mathfrak{C}_p, \\ 0, & \text{otherwise} \end{cases}$$

is well defined.

Step 2: $V\big(x(t)\big) \to 0$ as $t \to \infty$.
For $p \in \{1,\dots,N\}$, let $P_p, Q_p \in \mathbb{C}^{n\times n}$ be the matrices as in Definition 4.3.3 corresponding to the DAE $E_p\dot{x} = A_p x$. Let furthermore

$$\lambda_p := \min_{x \in \mathfrak{C}_p\setminus\{0\}} \frac{x^T Q_p x}{V_p(x)} = \min_{\substack{x\in\mathfrak{C}_p \\ V_p(x)=1}} x^\top Q_p x > 0,$$

where positivity follows from positive definiteness of V_p and Q_p on $\mathfrak{C}_p$. Consider a solution $x : \mathbb{R} \to \mathbb{R}^n$ of (4.1.1), then from Lemma 4.2.6 it follows that on each open interval (s,t) which does not contain a switching time of σ the function x is smooth and a local solution of $E_p\dot{x} = A_p x$, where $p = \sigma(\tau)$, $\tau \in (s,t)$. From $x(\tau) \in \mathfrak{C}_p$ for all $\tau \in (s,t)$ it follows that $V(x(\tau)) = V_p(x(\tau))$ for all $\tau \in (s,t)$ and

$$\tfrac{\mathrm{d}}{\mathrm{d}t} V_p\big(x(\tau)\big) = x(\tau)^\top Q_p x(\tau) \leq -\lambda_p V_p\big(x(\tau)\big).$$

Let $t \in \mathbb{R}$ be a jump of σ, then $x(t) = \Pi_{\sigma(t)} x(t-)$ and $x(t-) \in \mathfrak{C}_{\sigma(t-)}$, hence, by (4.3.3),

$$\begin{aligned} V(x(t)) &= V_{\sigma(t)}\big(x(t)\big) = V_{\sigma(t)}\big(\Pi_{\sigma(t)} x(t-)\big) \\ &\leq V_{\sigma(t-)}\big(x(t-)\big) = V\big(x(t-)\big) \end{aligned}$$

For $\lambda := \min_p \lambda_p$ it therefore follows

$$\forall t, t_0 \in \mathbb{R} \text{ with } t \geq t_0 : \quad V\big(x(t)\big) \leq e^{-\lambda(t-t_0)} V\big(x(t_0)\big),$$

which implies that $V\big(x(t)\big) \to 0$ for all solutions x of (4.1.1).

Step 3: Solutions tend to zero.
Seeking a contradiction, assume $x(t) \not\to 0$. Then there exists $\varepsilon > 0$ and a sequence $(s_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ with $s_i \to \infty$ as $i \to \infty$ such that $\|x(s_i)\| > \varepsilon$ for all $i \in \mathbb{N}$. There is at least one $p \in \{1, \ldots, N\}$ such that the set $\{\ i \in \mathbb{N} \mid \sigma(s_i) = p\ \}$ has infinitely many elements, therefore assume that $\sigma(s_i) = p$ for some p and all $i \in \mathbb{N}$. Then $x(s_i) \in \mathfrak{C}_p \setminus \{\ \xi \in \mathfrak{C}_p \mid \|\xi\| < \varepsilon\ \}$ for all $i \in \mathbb{N}$ and since V_p is positive definite on $\mathfrak{C}_p$ there exists $\delta > 0$ such that $V\big(x(s_i)\big) > \delta$ for all $i \in \mathbb{N}$. This is a contradiction to $V\big(x(t)\big) \to 0$ as $t \to \infty$. Therefore $x(t) \to 0$ as $t \to \infty$. $\boxed{\text{qed}}$

Remark 4.3.10 (Alternative condition)
Condition (4.3.3) is equivalent to the condition

$$\forall\, p, q \in \{1, \ldots, N\}\ \forall x \in \mathbb{R}^n : \quad V_p(\Pi_p \Pi_q x) \leq V_q(\Pi_q x)$$

which might be easier to check. □

Condition (4.3.3) implies that any two Lyapunov functions V_p and V_q coincide on the intersection $\mathfrak{C}_p \cap \mathfrak{C}_q$, hence Theorem 4.3.9 is a generalization of the switched ODE case where the existence of a common Lyapunov function is sufficient to ensure stability under arbitrary switching. However, the existence of a common Lyapunov function is not enough in the DAE case, as becomes clear from Example 1a in Section 4.3.2. Under arbitrary switching, solutions will in general exhibit jumps; these jumps are described by the consistency projectors, and these projectors must "fit together" with the Lyapunov functions in the sense of (4.3.3) to ensure stability of the switched DAE under arbitrary switching.

In Example 1b the projectors do not fit together with the Lyapunov function $V(x) = x^\top x$, but for the common Lyapunov function

$$V(x_1, x_2) = {x_1}^2 - x_1 x_2 + {x_2}^2$$

the condition (4.3.3) is fulfilled.

If one assumes that the switching signal is chosen in such a way that no jumps occur, then the conditions on the consistency projectors are not needed and the following corollary holds.

Corollary 4.3.11
Consider the switched DAE (4.1.1) satisfying (S1), (S2), (A2) and assume each DAE $E_p\dot{x} = A_p x$, $p = 1, \dots, N$, is asymptotically stable with Lyapunov function V_p and consistency spaces $\mathfrak{C}_p$. Let, for $t_0 \in \mathbb{R}$ and $x^0 \in \mathbb{R}^n$,

$$\Sigma_{(t_0,x^0)} := \left\{ \sigma : \mathbb{R} \to \{1, \dots, N\} \;\middle|\; \begin{array}{l} \sigma \text{ fulfills (A1) and } \exists \text{ solution } x \\ \text{of (4.1.1) with } x(t_0) = x^0 \\ \text{and } x \text{ has no jumps} \end{array} \right\}.$$

If

$$\forall\, p, q = 1, \dots, N \;\forall x \in \mathfrak{C}_p \cap \mathfrak{C}_q : \quad V_p(x) = V_q(x), \tag{4.3.4}$$

then, for $\sigma \in \Sigma_{(t_0,x^0)}$, $x^0 \in \mathbb{R}^n$, all solutions x of (4.1.1) with $x(t_0) = x^0 \in \mathbb{R}^n$ converge to zero as $t \to \infty$. □

Note that actually it is not necessary to impose Assumptions (S1), (S2) and (A2) explicitly in the above corollary: (S1) is already induced by the definition of $\Sigma_{(t_0,x^0)}$, (S2) follows from the assumption that each DAE $E_p\dot{x} = A_p x$ is asymptotically stable [Ber08], and (A2) is not needed any more, because the assumption that no jumps occur also implies that no impulses can occur, see Remark 4.2.12.

Example 3 from Section 4.3.2 fulfills the assumptions of Corollary 4.3.11, hence if no jumps occur all solutions tend to zero. In fact, the corresponding Lyapunov functions are

$$\begin{aligned} V_1(x) &= {x_1}^2 + {x_2}^2, \\ V_2(x) &= {x_2}^2 + 4{x_3}^2, \end{aligned}$$

which clearly coincide on the intersection $\mathfrak{C}_1 \cap \mathfrak{C}_2$ which is the x_2-axis. For every $x^0 = (x_1^0, x_2^0, x_3^0)^\top \in \mathbb{R}^n$ with $x_1^0 \neq 0$ and $x_3^0 \neq 0$ it follows that $x^0 \notin \mathfrak{C}_1 \cup \mathfrak{C}_2$ hence no jump free solution with $x(t_0) = x^0$ exists, i.e. $\Sigma_{(t_0,x^0)} = \emptyset$. If $x^0 \in \mathfrak{C}_1$ or $x^0 \in \mathfrak{C}_2$, then there clearly exists a unique minimal $\tau_{(t_0,x_0)}$ such that any solution x with $x(t_0) = x^0$ fulfills $x(t_0 + \tau_{(t_0,x^0)}) \in \mathfrak{C}_1 \cap \mathfrak{C}_2$ and

$$\Sigma_{(t_0,x^0)} = \left\{ \sigma : \mathbb{R} \to \{1,\ldots,N\} \;\middle|\; \begin{array}{l} \forall \text{switching times } t \text{ of } \sigma\ \exists k \in \mathbb{Z} : \\ t = t_0 + \tau_{(t_0,x^0)} + \frac{k}{2} \end{array} \right\}.$$

In contrast to this, in Example 1a and 1b the condition (4.3.4) is trivially fulfilled because $\mathfrak{C}_1 \cap \mathfrak{C}_2 = \{0\}$, but only the constant switching signals yield non-jumping non-trivial solutions, i.e. the set $\Sigma_{(t_0,x^0)}$ is practically always empty. Hence Corollary 4.3.11 is not very useful in this case. For Example 2 it is not possible to find Lyapunov functions for both subsystems such that condition (4.3.4) is fulfilled.

For switched ODEs it is well known that switching between stable subsystems always yields a stable system provided the so-called dwell time is large enough. Consider therefore the following set of switching signals parametrized by a dwell time $\tau_d > 0$:

$$\Sigma^{\tau_d} := \left\{ \sigma : \mathbb{R} \to \{1,\ldots,N\} \;\middle|\; \begin{array}{l} \forall \text{ switching times } t_i \in \mathbb{R}, i \in \mathbb{Z} : \\ t_{i+1} - t_i \geq \tau_d \end{array} \right\}.$$

Theorem 4.3.12 (Large dwell time ensures stability)
Consider the switched DAE (4.1.1) satisfying (S1), (S2), (A2) and assume that each DAE $E_p\dot{x} = A_p x$, $p = 1,\ldots,N$, is asymptotically stable with Lyapunov functions V_p and corresponding matrices $Q_p \in \mathbb{C}^{n\times n}$. Let

$$\lambda := \min_p \min_{x \in \mathfrak{C}_p \setminus \{0\}} \frac{x^T Q_p x}{V_p(x)}.$$

Let $\mu \geq 1$ be such that

$$\forall p,q = 1,\ldots,N\ \forall x \in \mathfrak{C}_q : \quad V_p(\Pi_p x) \leq \mu V_q(x). \tag{4.3.5}$$

Then the switched DAE (4.1.1) with $\sigma \in \Sigma^{\tau_d}$ is asymptotically stable whenever

$$\tau_d > \frac{\ln \mu}{\lambda}.$$

□

Proof. First note that all solutions of (4.1.1) are impulse free by Theorem 4.2.13. Fix a solution $x \in \mathbb{R} \to \mathbb{R}^n$ of (4.1.1) with a fixed switching signal $\sigma \in \Sigma^{\tau_d}$. If σ has only finitely many switching times, then asymptotic stability of (4.1.1) is obvious, therefore assume that the set of switching times $\{ \, t_i \in \mathbb{R} \mid i \in \mathbb{Z} \, \}$ of σ is infinite.

Let $v : \mathbb{R} \to \mathbb{R}_{\geq 0}$, $t \mapsto V_{\sigma(t)}\big(x(t)\big)$ and $0 < \varepsilon := \tau_d - \frac{\ln \mu}{\lambda}$. Then as in the proof of Theorem 4.3.9 it follows that

$$v(t_{i+1}-) \leq e^{-\lambda(t_{i+1}-t_i)} v(t_i) \leq \frac{e^{-\lambda\varepsilon}}{\mu} v(t_i).$$

Furthermore, condition (4.3.5) yields

$$v(t_i) = V_{\sigma(t_i)}\big(\Pi_{\sigma(t_i)} x(t_i-)\big) \leq \mu V_{\sigma(t_i-)}\big(x(t_i-)\big) = \mu v(t_i-).$$

All together this yields for all $i \in \mathbb{Z}$,

$$v(t_{i+1}-) \leq e^{-\lambda\varepsilon} v(t_i-),$$

hence $v(t_i-) \to 0$ as $i \to \infty$. Since $v(t) \leq e^{-\lambda(t-t_i)} v(t_i) \leq \mu v(t_i-)$ for all $t \in [t_i, t_{i+1})$, $i \in \mathbb{Z}$, it also follows that $v(t) \to 0$ as $t \to \infty$. As in the proof of Theorem 4.3.9 it now follows that $x(t) \to 0$ as $t \to 0$. qed

Remark 4.3.13 (Large enough dwell time always possible)
Since each Lyapunov function V_q, $q \in \{1, \ldots, N\}$, is a quadratic function which is positive definite on $\mathfrak{C}_q$, it follows that for all $p, q \in \{1, \ldots, N\}$

$$\mu_{q,p} := \min_{x \in \mathfrak{C}_q \setminus \{0\}} \frac{V_p\big(\Pi_p x\big)}{V_q(x)} = \min_{x \in \mathfrak{C}_q : V_q(x)=1} V_p\big(\Pi_p x\big) > 0,$$

hence (4.3.5) is always fulfilled for $\mu \geq \max_{q,p} \mu_{q,p}$. Therefore, Theorem 4.3.12 states that switching between asymptotically stable subsystems yields asymptotic stability provided the dwell time of the switching signal is large enough. □

For Example 1a from Section 4.3.2, condition (4.3.5) is fulfilled for the (common) Lyapunov function $x \mapsto V(x) = x^\top x$ with $\lambda = 2$ and $\mu = 2$. Hence for dwell times larger than $\frac{1}{2} \ln 2$ the switched system (4.1.1) is asymptotically stable, see also Figure 4.

5 Controllability and observability for distributional DAEs

5.1 Controllability

For classical DAEs there exist the two concepts of controllability: R-controllability and impulse controllability which in some sense are complementary [Dai89]. While R-controllability considers the reachability of any consistent state in some finite time, impulse-controllability is related to the reachability of certain impulsive parts. The latter controllability is based on instantaneous control, i.e. the value of the input and its derivatives $u^{(i)}(t+)$, $i \in \mathbb{N}$, determines the impulsive part $x[t]$. Although R-controllability is not defined as instantaneous control, it is well known that the "control interval" can be chosen arbitrarily small. In the limit, an instantaneous control results which is impulsive, so it is natural to generalize R-controllability as the ability to choose $u[t]$ such that all consistent values and derivatives $x^{(i)}(t+)$, $i \in \mathbb{N}$, can be reached. Since this controls the jump from $x^{(i)}(t-)$ to $x^{(i)}(t+)$ the term "jump-controllability" will be used in the following.

For classical time-varying systems, controllability and reachability are two different concept, in the first concept the control is applied in the future, while in the second concept the control is applied in the past. However, by shrinking the control-interval to length zero both concepts get indistinguishable. The same arguments yield that the two different concepts initial observability and final observability get conceptually identically when the observed interval shrinks to length zero.

For notational convenience let, for $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ and $t \in \mathbb{R}$,

$$D^{(*)}(t+) := \big(D(t+), D'(t+), \ldots, D^{(i)}(t+), \ldots\big) \in \mathbb{R}^{\mathbb{N}},$$

i.e. $D^{(*)}(t+)$ is the sequence of all right-sided derivatives.

Definition 5.1.1 (Jump- and impulse-controllability)
Consider a distributional DAE with an input

$$E\dot{x} = Ax + Bu \tag{5.1.1}$$

where $E, A \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{m\times n}$, $B \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times p}$, $m,n,p \in \mathbb{N}$. Let $\mathcal{B}^{t_0}_{(E,A,B)}$ be the set of all ITP solutions of (5.1.1) with initial time $t_0 \in \mathbb{R}$, i.e.

$$\mathcal{B}^{t_0}_{(E,A,B)} := \left\{\ (x,u) \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n+p}\ \middle|\ (E\dot{x})_{[t_0,\infty)} = (Ax+Bu)_{[t_0,\infty)}\ \right\},$$

then (5.1.1) is called *jump-controllable at* $t_0 \in \mathbb{R}$ if, and only if,

$$\begin{aligned}&\forall (x^0,u^0),(x^1,u^1) \in \mathcal{B}^{t_0}_{(E,A,B)}\ \exists (x,u) \in \mathcal{B}^{t_0}_{(E,A,B)} :\\ &(x,u)_{(-\infty,t_0)} = (x^0,u^0)_{(-\infty,t_0)}\ \wedge\ (x,u)^{(*)}(t_0+) = (x^1,u^1)^{(*)}(t_0+).\end{aligned}$$

The distributional DAE (5.1.1) is called *impulse-controllable at* $t_0 \in \mathbb{R}$ if, and only if,

$$\begin{aligned}&\forall (x^0,u^0),(x^1,u^1) \in \mathcal{B}^{t_0}_{(E,A,B)}\ \exists (x,u) \in \mathcal{B}^{t_0}_{(E,A,B)} :\\ &\qquad (x,u)_{(-\infty,t_0)} = (x^0,u^0)_{(-\infty,t_0)}\ \wedge\ (x,u)[t_0] = (x^1,u^1)[t_0].\end{aligned}$$ □

Remark 5.1.2 (Distributional behaviours)
For the definition of jump- and impulse-controllability, it is not assumed that the matrix pair (E,A) in (5.1.1) is DAE-regular. In fact, the definition can be easily generalized to *distributional behaviours*

$$\mathcal{B} = \left\{\ w \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^q\ \middle|\ R\big(\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\big)(w) = 0\ \right\},$$

where $R(\partial) \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{r\times q}[\partial]$ is a matrix polynomial, and the corresponding *initial trajectory distributional behaviour*

$$\mathcal{B}_{t_0} = \left\{\ w \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^q\ \middle|\ R\big(\tfrac{\mathrm{d}_\mathbb{D}}{\mathrm{d}t}\big)(w)_{[t_0,\infty)} = 0\ \right\}.$$

The generalization of the behavioural approach as introduced by Willems (see e.g. [Wil07] for an survey) to distributional behaviours seems an interesting research topic, in particular for the study of DAEs which are not DAE-regular. □

Let $\int_{t_0} : \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \to \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ be the antiderivative operator as in Proposition 2.1.18, then

$$v[t_0] = w[t_0] \quad \Leftrightarrow \quad \forall i \in \mathbb{N}_{>0} : \left(\left(\int_0\right)^i v\right)(t_0+) = \left(\left(\int_0\right)^i w\right)(t_0+),$$

hence jump- and impulse-observability are complementary in the following sense: jump-controllability is defined as the ability to match *all derivatives* of any trajectories while impulse-controllability is defined to match *all antiderivatives* of any trajectory.

The next proposition shows that it suffices to consider "controllability to zero".

Proposition 5.1.3 (Controllability to zero)
The distributional DAE (5.1.1) is jump-controllable at $t_0 \in \mathbb{R}$ if, and only if,

$$\forall (x^0, u^0) \in \mathcal{B}^{t_0}_{(E,A,B)}\ \exists (x,u) \in \mathcal{B}^{t_0}_{(E,A,B)} : \\ (x,u)_{(-\infty,t_0)} = (x^0,u^0)_{(-\infty,t_0)}\ \wedge\ (x,u)^{(*)}(t_0+) = 0 \quad (5.1.2)$$

and (5.1.1) is impulse-controllable at t_0 if, and only if,

$$\forall (x,u)^0 \in \mathcal{B}_{t_0}\ \exists (x,u) \in \mathcal{B}^{t_0}_{(E,A,B)} : \\ (x,u)_{(-\infty,t_0)} = (x^0,u^0)_{(-\infty,t_0)}\ \wedge\ (x,u)[t_0] = 0. \quad (5.1.3)$$

$\square$

Proof. For notational convenience write $w := (x,u)$.

Let (5.1.1) be controllable at t_0 and let $w^0 \in \mathcal{B}^{t_0}_{(E,A,B)}$ be arbitrary. For $w^1 := 2w^0 \in \mathcal{B}^{t_0}_{(E,A,B)}$ and $w^2 := w^0$ choose $\overline{w} \in \mathcal{B}^{t_0}_{(E,A,B)}$ such that $\overline{w}_{(-\infty,t_0)} = {w^1}_{(-\infty,t_0)}$ and $\overline{w}^{(*)}(t_0+) = (w^2)^{(*)}(t_0+)$. Then $w := \overline{w} - w^0 \in \mathcal{B}^{t_0}_{(E,A,B)}$ fulfills $w_{(-\infty,t_0)} = (w^1 - w^0)_{(-\infty,t_0)} = {w^0}_{(-\infty,t_0)}$ and $w^{(*)}(t_0+) = (w^2 - w^0)^{(*)}(t_0+) = 0$. Hence "jump-controllability to zero" is shown.

Assume now that (5.1.2) holds and let $w^1, w^2 \in \mathcal{B}^{t_0}_{(E,A,B)}$ be arbitrary. For $w^0 := w^1 - w^2$ choose $\overline{w} \in \mathcal{B}^{t_0}_{(E,A,B)}$ such that $\overline{w}_{(-\infty,t_0)} = {w^0}_{(-\infty,t_0)}$ and $\overline{w}^{(*)}(t_0+) = 0$. Then $w := \overline{w} + w^2 \in \mathcal{B}^{t_0}_{(E,A,B)}$ fulfills $w_{(\infty,t_0)} = (\overline{w} + w^2)_{(-\infty,t_0)} = {w^1}_{(-\infty,t_0)}$ and $w^{(*)}(t_0+) = (\overline{w} + w^2)^{(*)}(t_0+) = (w^2)^{(*)}(t_0+)$, hence (5.1.1) is controllable at t_0.

An analogous argumentation shows that (5.1.1) is impulse-controllable if, and only if, (5.1.3) holds. $\boxed{\text{qed}}$

The following results show that the above defined controllability concepts generalize the well known controllability concepts of classical regular DAEs $E\dot{x} = Ax + Bu$ with constant matrices $E, A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$.

Proposition 5.1.4 (Generalization of classical controllability)
For $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $t_0 \in \mathbb{R}$ and a nilpotent matrix $N \in \mathbb{R}^{n\times n}$ the following holds.

(i) The (distributional) ODE $\dot{x} = Ax + Bu$ is jump-controllable (at any t_0) if, and only if, $\operatorname{rk}[B, AB, \ldots, A^{n-1}B] = n$, i.e. the classical ODE $\dot{x} = Ax + Bu$ is controllable [Kal60].

(ii) The (distributional) ODE $\dot{x} = Ax + Bu$ is always impulse-controllable.

(iii) The (distributional) pure DAE $N\dot{x} = x + Bu$ is always jump-controllable.

(iv) The (distributional) pure DAE $N\dot{x} = x + Bu$ is impulse-controllable (at any t_0) if, and only if, $\operatorname{im} N = \operatorname{im}[NB, N^2B, \ldots, N^{n-1}B]$, i.e. the classical pure DAE $N\dot{x} = x + Bu$ is impulse controllable in the classical sense ([Cob84, Dai89]). □

Proof. In the following, the characterizations (5.1.2) and (5.1.3) will be used.

First observe that in all cases the condition $(x^0, u^0) \in \mathcal{B}^{t_0}_{(E,A,B)}$ puts no restriction on $(x^0, u^0)_{(-\infty,t_0)}$ because the matrix pairs (I, A) and (N, I) are DAE-regular.

(i) For each $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ and $u \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$ there exists a unique solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of the corresponding ITP. In particular, by evaluating the impulsive part of the equation $\dot{x} = Ax + Bu$ at t_0,
$$\big(x(t_0+) - x^0(t_0-)\big)\delta_{t_0} + x[t_0]' = Ax[t_0] + Bu[t_0].$$
The ansatz $x[t_0] = \sum_{i=0}^k \alpha_i \delta^i_{t_0}$ and $u[t_0] = \sum_{j=0}^l \beta_i \delta^i_{t_0}$ for some $k, l \in \mathbb{R}$ with $k < l$ and $\alpha_0, \ldots, \alpha_k \in \mathbb{R}^n$, $\beta_0, \ldots, \beta_l \in \mathbb{R}^m$ yields

$x[t_0]' = \sum_{i=0}^{k} \alpha_i \delta_{t_0}^{i+1}$ and therefore

$$\begin{aligned}
\delta_{t_0}: &\quad x(t_0+) - x^0(t_0-) = A\alpha_0 + B\beta_0\\
\delta_{t_0}': &\quad \alpha_0 = A\alpha_1 + B\beta_1\\
&\quad \vdots\\
\delta_{t_0}^{(k)}: &\quad \alpha_{k-1} = A\alpha_k + B\beta_k\\
\delta_{t_0}^{(k+1)}: &\quad \alpha_k = B\beta_{k+1}\\
\delta_{t_0}^{(k+2)}, \ldots, \delta_{t_0}^{(l)}: &\quad 0 = B\beta_{k+2} = \ldots = B\beta_l
\end{aligned}$$

If $\dot{x} = Ax + Bu$ is jump-controllable, then there exists $k \in \mathbb{N}$ and $\beta_0, \ldots, \beta_{k+1}$ such that $x(t_0+) = 0$ and

$$-x^0(t_0-) = [B, AB, \ldots, A^{k+1}B] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{k+1} \end{bmatrix}.$$

Since $x^0(t_0-) \in \mathbb{R}^n$ can be arbitrary, jump-controllability implies that there exists $k \in \mathbb{N}$ such that $\operatorname{rk}[B, AB, \ldots, A^{k+1}B] = n$, by the Cayley-Hamilton-Theorem the latter is equivalent to $\operatorname{rk}[B, AB, \ldots, A^{n-1}B] = n$.

If $\operatorname{rk}[B, AB, \ldots, A^{n-1}B] = n$, then $u[t_0]$ can be chosen such that the above ansatz yields $\big(x(t_0+) - x^0(t_0-)\big) = -x^0(t_0-)$ which implies $x(t_0+) = 0$. Furthermore, (3.3.4) together with $u_{(t_0,\infty)} = 0$ implies that $x_{(t_0,\infty)} = 0$ and, in particular, $x^{(i)}(t_0+) = 0$ for all $i \in \mathbb{N}$, hence $\mathcal{B}_{(I,A,B)}$ is jump-controllable.

(ii) It must be shown that for every input $u \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$ with $u[t_0] = 0$, $t_0 \in \mathbb{R}$, the solution $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ of the ITP $\dot{x} = Ax + Bu$ with arbitrary initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ and initial time $t_0 \in \mathbb{R}$ fulfills $x[t_0] = 0$. Making again the ansatz $x[t_0] = \sum_{i=0}^{k} \alpha_i \delta_{t_0}^{(i)}$ for some $k \in \mathbb{N}$ and $\alpha_0, \ldots, \alpha_k \in \mathbb{R}^n$, the equation $\dot{x}[t_0] = (Ax +$

$Bu)[t_0]$ implies

$$x^0(t_0+) - x^0(t_0-) = A\alpha_0,\ \alpha_0 = A\alpha_1,\ \ldots,\ \alpha_{k-1} = A\alpha_k,\ \alpha_k = 0,$$

hence $0 = \alpha_k = \alpha_{k-1} = \ldots = \alpha_1 = \alpha_0$, i.e. $x[t_0] = 0$.

(iii) In Lemma 4.2.6 it was already shown that the local solution of $N\dot{x} = x + Bu$ on (t_0, ∞) with $u_{(t_0,\infty)} = 0$ fulfills $x_{(t_0,\infty)} = 0$, hence (5.1.2) holds.

(iv) Let $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be some arbitrary initial trajectory and let $u \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$ be some input with $u[t_0] = 0$. By Corollary 3.4.4 the unique solution of the corresponding ITP is explicitly given by

$$x = -\sum_{i=0}^{n-1} (N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^i (Bu_{(t_0,\infty)} - x^0_{(-\infty,t_0)}).$$

First observe that, for $i = 1, \ldots, n-1$,

$$\begin{aligned}(N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^i (Bu_{(t_0,\infty)} - x^0_{(-\infty,t_0)})\\ = N^i x^0(t_0-)\delta_{t_0}^{(i-1)} + N^i B u^{(i)}_{(t_0,\infty)} + N^i B \sum_{j=0}^{i-1} u^{(i-1-j)} \delta_{t_0}^{(j)},\end{aligned}$$

hence

$$x[t_0] = -\sum_{i=1}^{n-1} N^i \left(x^0(t_0-) + \sum_{j=0}^{n-2} N^j B u^{(j)}(t_0+) \right) \delta_{t_0}^{(i-1)}.$$

A necessary condition for $x[t_0] = 0$ is that at least the coefficient of δ_{t_0} is zero, i.e.

$$\forall x_0 \in \mathbb{R}^n\ \exists u_0, \ldots, u_{n-2} \in \mathbb{R}^m : \quad 0 = Nx_0 + N \sum_{j=0}^{n-2} N^j u_j,$$

or, in other words, $\operatorname{im} N = \operatorname{im}[NB, N^2B, \ldots, N^{n-1}B]$.

This condition is also sufficient for impulse controllability because choosing u such that $u^{(j)}(t_0+) = u_j$ as above yields $x[t_0] = 0$.

qed

Corollary 5.1.5
The classical DAE $E\dot{x} = Ax + Bu$ with regular matrix pair $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ is R-controllable (in the sense of [Dai89]) if, and only if, the corresponding distributional DAE 5.1.1 is jump-controllable (at any $t_0 \in \mathbb{R}$); it is impulse controllable (in the sense of [Dai89] or [Cob84]) if, and only if, (5.1.1) is impulse-controllable (at any $t_0 \in \mathbb{R}$). $\square$

Proof. Clearly, $E\dot{x} = Ax + Bu$ is jump- or impulse-controllable if, and only if, $SET\dot{x} = SATx + SBu$ for some constant invertible matrices $S, T \in \mathbb{R}^{n\times n}$ is jump- or impulse-controllable, respectively. Now the claim follows immediately from [Dai89, Thm. 2-2.2] and [Dai89, Thm. 2-2.3] $\boxed{\text{qed}}$

5.2 Observability

Consider now (5.1.1) with an output, i.e.

$$\begin{aligned} E\dot{x} &= Ax + Bu, \\ y &= Cx + Du, \end{aligned} \tag{5.2.1}$$

where $C \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{q\times n}$ and $D \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{q\times p}$, $q \in \mathbb{N}$.

Both, the input and the output, can be viewed as *external* signals while the variable x is an internal signal. It is an important property of a system as to whether it is possible to deduce the internal signals from the knowledge of the external signals.

For classical linear ODEs, observability is defined just in this way where it is assumed that the external signals are known on some interval. It is well known that for this definition of observability the length of the "observed" interval can be arbitrarily small, therefore, analogously as in the controllability definition, it is natural to only consider the family of all derivatives of the external signal at some time t_0 and check whether it is possible to deduce the internal signals from this information, this motivates the forthcoming definition of "jump-observability". For distributional DAEs the external and internal signals also exhibit impulsive parts, the observability of those parts is captured by the definition of "impulse-observability".

Definition 5.2.1 (Jump-observability and impulse-observability)
For (5.2.1) and $t_0 \in \mathbb{R}$ let

$$\mathcal{B}^{t_0}_{(E,A,B,C,D)} := \left\{ (x,u,y) \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}) \;\middle|\; \begin{array}{r} (E\dot{x})_{[t_0,\infty)} = (Ax+Bu)_{[t_0,\infty)}, \\ y_{[t_0,\infty)} = (Cx+Du)_{[t_0,\infty)} \end{array} \right\}.$$

The distributional DAE (5.2.1) is called *jump-observable* at t_0 if, and only if,

$$\begin{array}{l} \forall (x^0,u^0,y^0),(x^1,u^1,y^1) \in \mathcal{B}^{t_0}_{(E,A,B,C,D)} : \\ (u^0,y^0)^{(*)}(t_0+) = (u^1,y^1)^{(*)}(t_0+) \quad \Rightarrow \quad (x^0)^{(*)}(t_0+) = (x^1)^{(*)}(t_0+). \end{array}$$

The distributional DAE (5.2.1) is called *impulse-observable* at t_0 if, and only if,

$$\begin{array}{l} \forall (x^0,u^0,y^0),(x^1,u^1,y^1) \in \mathcal{B}^{t_0}_{(E,A,B,C,D)} : \\ \qquad (u^0,y^0)[t_0] = (u^1,y^1)[t_0] \quad \Rightarrow \quad x^0[t_0] = x^1[t_0]. \end{array}$$ □

Remark 5.2.2 (The feed-through term Du)
The feed-through term Du in (5.2.1) is in most situation not needed explicitly because it can be incorporated into the system by adding an additional state variable z and the algebraic condition $0 = z - Du$ and rewriting the output equation as $y = [C, I](x/z)$. If $D[\cdot] = 0$, then it is easy to see that this will not change the observability and controllability properties. In the general case, i.e. $D[t_0] \neq 0$ for some $t_0 \in \mathbb{R}$, the impulse-controllability-property might be changed, because then it is possible that $z[t_0] \neq 0$ independently of $u^{(*)}(t_0+)$ and although $u[t_0] = 0$. □

Remark 5.2.3 (Distributional behaviours)
Remark 5.1.2 for distributional behaviours also applies for the observability definition. However there is a significant difference: (jump- or impulse-)controllability is a property of the behaviour itself, it is not necessary to explicitly define an input, whilst for observability it is necessary to first define "external signals" from which the internal signals should be observed. Formally, for a given behaviour

$\mathcal{B} = \ker R(\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}) \subseteq (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^q$ as in Remark 5.1.2 let the "external behaviour" be given as

$$\begin{aligned}\mathcal{B}_{\mathrm{ext}} &= \left\{ \; v \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m \;\middle|\; \exists w \in \mathcal{B} : \; v = M(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})(w) \; \right\} \\ &= M(\tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})\mathcal{B} \subseteq (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m,\end{aligned}$$

where $M(\partial) \in \big(\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}[\partial]\big)^{m\times q}$ is some polynomial matrix.

For a distributional DAE (5.2.1) let the behaviour $\mathcal{B}$ be the set of all trajectories (x,u), then the corresponding external behaviour consists of all trajectories (u,y), i.e. $\mathcal{B}_{\mathrm{ext}} = \left[\begin{smallmatrix} 0 & I \\ C & D\end{smallmatrix}\right] \mathcal{B}$. □

Similar as for controllability it is first shown that observability is characterized by observability of zero.

Proposition 5.2.4 (Observability of zero)
The distributional DAE (5.2.1) is impulse-observable at t_0 if, and only if,

$$\forall (x,u,y) \in \mathcal{B}^{t_0}_{(E,A,B,C,D)} : \quad (u,y)^{(*)}(t_0+) = 0 \;\Rightarrow\; x^{(*)}(t_0+) = 0. \tag{5.2.2}$$

The distributional DAE (5.2.1) is impulse-observable at t_0 if, and only if,

$$\forall (x,u,y) \in \mathcal{B}^{t_0}_{(E,A,B,C,D)} : \quad (u,y)[t_0] = 0 \;\Rightarrow\; w[t_0] = 0. \tag{5.2.3}$$

□

Proof. If (5.2.1) is jump-observable (or impulse-observable) at t_0 the property (5.2.2) (or (5.2.3)) follows easily by considering $(x^1,u^1,y^1) = 0$. For the converse just consider the difference $(x^0 - x^1, u^0 - u^1, y^0 - y^1)$. qed

For the next results, classical DAEs with inputs and outputs are considered:

$$\begin{aligned} E\dot{x} &= Ax + Bu, \\ y &= Cx, \end{aligned}$$

where $E, A \in \mathbb{R}^{n\times n}$ are such that the matrix pair (E, A) is regular, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{k\times n}$. As mentioned in Remark 5.2.2 it is not a restriction of the general case to set $D = 0$ in (5.2.1) for the constant coefficient case.

Proposition 5.2.5 (Generalization of classical observability)
For $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{k\times n}$, $t_0 \in \mathbb{R}$ and a nilpotent matrix $N \in \mathbb{R}^{n\times n}$ the following holds.

(i) The (distributional) ODE $\dot{x} = Ax + Bu$, $y = Cx$, is jump-observable (at any t_0) if, and only if, $\operatorname{rk}[C/CA/\cdots/CA^{n-1}] = n$, i.e. the classical ODE $\dot{x} = Ax+Bu$, $y = Cx$ is observable [Kal60].

(ii) The (distributional) ODE $\dot{x} = Ax + Bu$, $y = Cx$, is always impulse-observable.

(iii) The (distributional) pure DAE $N\dot{x} = x + Bu$, $y = Cx$, is always jump-observable.

(iv) The (distributional) pure DAE $N\dot{x} = x+Bu$, $y = Cx$, is impulse-observable if, and only if, $\ker N = \ker[CN, CN^2, \ldots, CN^{n-1}]$, i.e. the classical pure DAE $N\dot{x} = x + Bu$, $y = Cx$ is impulse observable in the classical sense ([Cob84, Dai89]).

Proof. In the following the characterizations (5.2.2) and (5.2.3) for jump- and impulse-observability will be used.

(i) For the initial trajectory $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ let $x \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be the unique solution of the corresponding ITP for $u = 0$. It then follows that

$$y(t_0+) = Cx(t_0+), \quad y'(t_0+) = Cx'(t_0+) = CAx(t_0+), \\ \ldots, \quad y^{(i)} = CA^i x(t_0+).$$

Corollary 3.3.9 yields that $x(t_0+) = x(t_0-) = x^0(t_0-)$ hence $x(t_0+) \in \mathbb{R}^n$ can be arbitrary. Hence, by the Cayleigh-Hamilton Theorem, jump-observability is equivalent to the full rank condition $\operatorname{rk}[C, CA, \ldots, CA^{n-1}] = n$.

(ii) As in the proof of Proposition 5.1.4, $u[t_0] = 0$ implies that all ITP solutions of $\dot{x} = Ax + Bu$ fulfill $x[t_0] = 0$, hence impulse-observability holds.

(iii) Let $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be any initial trajectory and let $u \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$ be some input signal. By Corollary 3.4.4 the unique solution of the corresponding ITP is explicitly given by

$$x = -\sum_{i=0}^{n-1} (N_{[t_0,\infty)} \tfrac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t})^i (Bu_{(t_0,\infty)} - x^0_{(-\infty,t_0)}),$$

therefore

$$x_{(t_0+,\infty)} = -\sum_{i=0}^{n-1} N^i B u^{(i)}_{(t_0,\infty)}.$$

In particular, $u^{(*)}(t_0+) = 0$ always implies $x^{(*)}(t_0+) = 0$, hence $\mathcal{B}_{(N,I,B)}$ is jump-observable (independently from the actual output).

(iv) Let $x^0 \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$ be some arbitrary initial trajectory and let $u \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^m$ be some input with $u[t_0] = 0$. As in the proof of Proposition 5.1.4 it follows that the unique solution x of the corresponding ITP fulfills

$$x[t_0] = -\sum_{i=1}^{n-1} N^i \left(x^0(t_0-) + \sum_{j=0}^{n-2} N^j B u^{(j)}(t_0+) \right) \delta_{t_0}^{(i-1)}$$

and

$$y[t_0] = Cx[t_0] = -\sum_{i=1}^{n-1} CN^i \left(x^0(t_0-) + \sum_{j=0}^{n-2} N^j B u^{(j)}(t_0+) \right) \delta_{t_0}^{(i-1)}.$$

Since $x(t_0-) \in \mathbb{R}^n$ is arbitrary, the implication $y[t_0] = 0 \Rightarrow x[t_0] = 0$ holds if, and only if,

$$\ker[CN/CN^2/\cdots/CN^{n-1}] = \ker[N/N^2/\cdots/N^{n-1}] = \ker N.$$

qed

Corollary 5.2.6
The classical DAE $E\dot{x} = Ax + Bu$, $y = Cx$, with regular matrix pair $(E, A) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$ is R-observable (in the sense of [Dai89]) if, and only if, the corresponding distributional behaviour $\mathcal{B}_{(E,A,B)}$ is jump-observable (at any $t_0 \in \mathbb{R}$) and it is impulse-observable (in the sense of [Dai89] or [Cob84]) if, and only if, $\mathcal{B}_{(E,A,B)}$ is impulse-observable (at any $t_0 \in \mathbb{R}$). $\square$

Proof. This follows immediately from [Dai89, Thm. 2-3.3] and [Dai89, Thm. 2-3.4]. qed

5.3 A normal form for pure DAEs

The previous sections showed that for regular DAEs with constant coefficients the controllability and observability properties are complementary in the sense that jump-controllability and -observability are properties solely given by the underlying ODE and impulse-controllability and -observability are properties of the underlying pure DAE. For ODEs there are several normal forms incorporating the input and output. The Byrnes-Isidori normal form (which focuses on the relative degree [Isi95, p. 165], see also [IRT07, Lem. 3.5]) and the Kalman-decomposition (which focuses on controllable and observable substates [Kal62]) are examples of such normal forms.

The aim of this section is to find similar normal forms for time-invariant pure DAEs. It will turn out that there exists a normal form which actually combines the properties of the above mentioned normal forms for ODEs. In fact, the state space is separated into impulse-controllable and -observable sub-states and, simultaneously, the so called relative degree determines the structure of the normal form, see Theorem 5.3.12. Compared to a similar decomposition proposed in [Dai89, p. 52] (without proof) the normal form from Theorem 5.3.10 is more specific and allows for a better analysis.

There are already results on normal or condensed forms of DAEs available, e.g. [VLK81], [LÖMK91], [Rat97], [KM06], but they do not focus on the relative degree or on impulse-controllable and -observable states. In addition, they partly use a different concept of equivalence which leads to other normal forms.

The considered equivalence relation used to obtain the normal form is given in the following.

Definition 5.3.1 (Equivalence of DAEs)
Identify each DAE

$$\begin{aligned} E\dot{x} &= Ax + bu, \\ y &= cx, \end{aligned} \tag{5.3.1}$$

with the corresponding tuple $(c, E, A, b) \in \mathbb{R}^{1\times n} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n} \times \mathbb{R}^m$ and *size* $m\times n$. Two DAEs (c_1, E_1, A_1, b_1) and (c_2, E_2, A_2, b_2) are called equivalent, or short

$$(c_1, E_1, A_1, b_1) \simeq (c_2, E_2, A_2, b_2),$$

if, and only if, they have the same size $m \times n$ and there exist invertible matrices $S \in \mathbb{R}^{m\times m}$, $T \in \mathbb{R}^{n\times n}$ such that

$$(c_2, E_2, A_2, b_2) = (c_1 T, SE_1 T, SA_1 T, Sb_1). \qquad \square$$

Remark 5.3.2 (Equivalence and solutions)
With the notation as in Definition 5.3.1 the following equivalence holds for all $(y, x, u) \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty} \times (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n \times \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$:

$$(y, x, u) \text{ solves } (c_1, E_1, A_1, b_1) \;\Leftrightarrow\; (y, T^{-1}x, u) \text{ solves } (c_2, E_2, A_2, b_2), \qquad \square$$

in particular, the inputs and outputs remain "unchanged". The same is true also for ITP problems, provided the initial trajectories x_0^1 and x_0^2 fulfill $x_0^1 = Tx_0^2$ (compare also Remark 3.1.6).

Definition 5.3.3 (Pure DAEs)
A DAE (c, E, A, b) with size $n \times n$ is called *pure DAE* if, and only if, there exists a nilpotent matrix $N \in \mathbb{R}^{n\times n}$ such that

$$(c, E, A, b) \simeq (\widehat{c}, N, I, \widehat{b}),$$

with corresponding $\widehat{c}^\top, \widehat{b} \in \mathbb{R}^n$. A pure DAE (c, E, A, b) is called in standard form if, and only if, $A = I$. $\square$

Remark 5.3.4 (Pure DAEs are pure distributional DAEs)
Clearly, (c, E, A, b) is a pure DAE if, and only if, A is invertible and $A^{-1}E$ (or, equivalently, EA^{-1}) is nilpotent. Hence the definition is consistent with Definition 3.4.1. □

From the theory of linear ODEs it is well known, that the controllability- and observability-matrices play an important role for controllability and observability as well as for the construction of normal forms. It is possible to define analogous matrices for DAEs, which play similar roles. Furthermore one can define impulse-controllability- and impulse-observability-indices which are invariants with respect to equivalence transformations. This is important for the normal form and can be used for characterizations of impulse-controllability and -observability.

Definition 5.3.5 (Impulse-controllability/observability-index)
Consider a pure DAE (c, E, A, b) with size $n \times n$.
The *impulse-controllability-matrix* of (c, E, A, b) is

$$\mathfrak{B}_{\text{imp}} := \left[b, N_b b, {N_b}^2 b, \ldots, {N_b}^{n-1} b\right], \quad \text{where } N_b := EA^{-1}.$$

The *impulse-controllability-index* of (c, E, A, b) is

$$d_b := \operatorname{rk} \mathfrak{B}_{\text{imp}}.$$

The *impulse-observability-matrix* of (c, E, A, b) is

$$\mathfrak{C}_{\text{imp}} := \left[c / cN_c / c{N_c}^2 / \ldots / c{N_c}^{n-1}\right], \quad \text{where } N_c := A^{-1}E.$$

The *impulse-observability-index* of (c, E, A, b) is

$$d_c := \operatorname{rk} \mathfrak{C}_{\text{imp}}.$$ □

Proposition 5.3.6 (Invariance)
Impulse-controllability and -observability as well as the corresponding indices are invariant under equivalence transformations. □

Proof. It follows from Remark 5.3.2 that an equivalence transformation does not change the property of a system to be impulse-controllable or -observable.

Let $(c_1, E_1, A_1, b_1) \simeq (c_2, E_2, A_2, b_2)$ via $S, T \in \mathbb{R}^{n\times n}$ and let $\mathfrak{B}_{\mathrm{imp},1}$, $\mathfrak{C}_{\mathrm{imp},1}$, $\mathfrak{B}_{\mathrm{imp},2}$, and $\mathfrak{C}_{\mathrm{imp},2}$ be the corresponding impulse-controllable- and impulse-observable-matrices. From the definition it follows that

$$\mathfrak{B}_{\mathrm{imp},2} = S\mathfrak{B}_{\mathrm{imp},1},$$

hence the corresponding impulse-controllability-indices are equal. Analogously,

$$\mathfrak{C}_{\mathrm{imp},2} = \mathfrak{C}_{\mathrm{imp},1}T$$

which shows that the impulse-observability-index is invariant. $\boxed{\text{qed}}$

Proposition 5.3.7 (Special structure of $\mathfrak{B}_{\mathrm{imp}}$ and $\mathfrak{C}_{\mathrm{imp}}$)
Consider a pure DAE in standard form (c, N, I, b) with size $n \times n$ and with impulse-controllability- and impulse-observability-indices $d_b, d_c \in \mathbb{N}$, respectively. Then

$$\mathfrak{B}_{\mathrm{imp}} = \big[b, Nb, \ldots, N^{d_b-1}b, 0, \ldots, 0\big]$$

and

$$\mathfrak{C}_{\mathrm{imp}} = \big[c/cN/\ldots/cN^{d_c-1}/0/\ldots/0\big].$$ □

Proof. Let $d \in \mathbb{N}$ be the smallest number such that $N^d b = 0$ (which exists since N is nilpotent). In terms of [Lan70, XII.7] the vector b is N-cyclic with period d. Now [Lan70, Lemma XII.7.1] states that $[b, Nb, \ldots, N^{d-1}b]$ has full rank which yields

$$d_b = \mathrm{rk}\mathfrak{B}_{\mathrm{imp}} = \mathrm{rk}\big[b, Nb, \ldots, N^{d-1}b, 0, \ldots, 0\big] = d,$$

this is the assertion of the proposition. The same argument applied to $N^\top$ and $c^\top$ shows the analogous property for $\mathfrak{C}_{\mathrm{imp}}$. $\boxed{\text{qed}}$

Definition 5.3.8 (Negative relative degree)
Consider a pure DAE (c, E, A, b) with corresponding standard form $(\widehat{c}, N, I, \widehat{b})$. The *negative relative degree* $r \in \mathbb{N}$ is given by

$$r = \max_{i\in\mathbb{N}}\{\widehat{c}N^i\widehat{b} \neq 0\}.$$

If $\widehat{c}N^i\widehat{b} = 0$ for all $i \in \mathbb{N}$, let $r := -\infty$. □

Note that the negative relative degree is invariant under equivalence transformations.

Remark 5.3.9 (Classical relative degree)
Let $g(s) := c(Es - A)^{-1}b \in \mathbb{R}(s)$ be the so called transfer function of the DAE (c, E, A, b); if (c, E, A, b) is an ODE, i.e. E is invertible, then the (classical) relative degree ϱ is defined as

$$\varrho = \deg q(s) - \deg p(s),$$

where $p, q \in \mathbb{R}[s]$ are such that $g(s) = p(s)/q(s)$. It is not difficult to see that for pure DAEs the transfer function $g(s)$ is a polynomial and

$$r = \deg g(s) = -\varrho,$$

therefore Definition 5.3.8 is consistent with the classical definition. □

It is now possible to formulate the main result of this section. With the proposed normal form, the influence of the input on the states and the influence of states on the output can easily be seen.

Theorem 5.3.10 (Normalform for pure DAEs)
Consider a pure DAE (c, E, A, b) with size n, negative relative degree $r \geq 0$, impulse-controllability- and impulse-observability-indices $d_b, d_c \in \mathbb{N}$, respectively.

Then (c, E, A, b) is equivalent to $(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b})$, where

$$
\widehat{c} = [0, \ldots, 0, 1], \quad \widehat{N} = \begin{array}{@{}c@{}l}
\left[\begin{array}{c|c|c|c}
\begin{smallmatrix}0 & & \\ 1 & \ddots & \\ & 1 & 0\end{smallmatrix} & 0 & 0 & 0 \\ \hline
E_1 & N_1 & 0 & 0 \\ \hline
E_2 & E_3 & \begin{smallmatrix}0 & & \\ 1 & \ddots & \\ & 1 & 0\end{smallmatrix} & 0^* \\ \hline
0_* & 0 & 0 & \begin{smallmatrix}0 & & \\ 1 & \ddots & \\ & 1 & 0\end{smallmatrix}
\end{array}\right]
&
\begin{array}{l}
\} \, d_c - r - 1 \\
\} \, n - d_c - d_b + r + 1 \\
\} \, d_b - r - 1 \\
\} \, r + 1
\end{array}
\end{array}
$$

$$
\widehat{I} = \left[\begin{array}{c|c|c|c}
I & & & \\ \hline
 & I & & \\ \hline
 & & I & \\ \hline
 & & & I_*
\end{array}\right], \quad \widehat{b} = \left[\begin{array}{c}
0 \\ \hline
0 \\ \hline
0 \\ \hline
\begin{smallmatrix}\gamma \\ 0 \\ \vdots \\ 0\end{smallmatrix}
\end{array}\right],
$$

where $\gamma := cA^{-1}(EA^{-1})^r b = c(A^{-1}E)^r A^{-1} b \neq 0$,

$$
0^* = \begin{bmatrix} * & \cdots & * & 1 \\ & 0 & & \end{bmatrix}, \quad 0_* = \begin{bmatrix} & 1 \\ & * \\ 0 & \vdots \\ & * \end{bmatrix}, \quad I_* = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ * & \cdots & * & 1 \end{bmatrix},
$$

and $N_1 \in \mathbb{R}^{(n-d_c-d_b+r+1)\times(n-d_c-d_b+r+1)}$ is a nilpotent matrix (in Jordan canonical form). □

Proof. Without loss of generality, assume that the DAE is in standard form, i.e. $(c, E, A, b) = (c, N, I, b)$ for some nilpotent matrix N. In this case $\gamma = cN^r b \neq 0$.

The proof consists of two main steps. The first step is the construction of the transformation matrices S and T, in particular the construction must ensure that S and T are invertible. In the second step it is shown that indeed $(c, N, I, b) \simeq (\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b})$ via S and T.

Step 1: Construction of S and T

The construction is based on the five matrices $\mathcal{L} \in \mathbb{R}^{n\times(d_c-r-1)}$, $\overline{\mathcal{L}} \in \mathbb{R}^{n\times(n-d_c-d_b+r+1)}$, $\overline{\mathcal{B}} \in \mathbb{R}^{n\times(d_b-r-1)}$, $\mathcal{B} \in \mathbb{R}^{n\times(r+1)}$, and $\widehat{I} \in \mathbb{R}^{n\times n}$, which define the transformation matrix S and T by

$$S := \gamma \left[\mathcal{L}, \overline{\mathcal{L}}, \overline{\mathcal{B}}, \mathcal{B}\right]^{-1},$$
$$T := \frac{1}{\gamma}\left[\mathcal{L}, \overline{\mathcal{L}}, \overline{\mathcal{B}}, \mathcal{B}\right] \widehat{I}.$$

Step 1a: The matrix $\widehat{I}$.

Let

$$\widehat{I} := \left[\begin{array}{c|c} I & \\ \hline & I_* \end{array}\right] \in \mathbb{R}^{n\times n},$$

where

$$I_* := \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ \frac{cb}{-\gamma} & \frac{cNb}{-\gamma} & \dots & \frac{cN^{r-1}b}{-\gamma} & 1 \end{bmatrix} \in \mathbb{R}^{(r+1)\times(r+1)}. \tag{5.3.2}$$

Obviously, $\widehat{I}$ is invertible.

Step 1b: The matrices $\mathcal{B}$ and $\overline{\mathcal{B}}$.

Let

$$\mathcal{B} := \left[b, Nb, \dots, N^r b\right] \in \mathbb{R}^{n\times(r+1)}$$

and

$$\mathcal{C} := \left[cN^r / \dots / cN / c\right] \in \mathbb{R}^{(r+1)\times n}$$

Since $cN^k b = 0$ for all $k > r$ and $cN^r b \neq 0$ it follows that

$$\mathcal{CB} = \begin{bmatrix} cN^r b & & & \\ cN^{r-1}b & \ddots & & \\ \vdots & \ddots & \ddots & \\ cb & \dots & cN^{r-1}b & cN^r b \end{bmatrix} \in \mathbb{R}^{(r+1)\times(r+1)} \tag{5.3.3}$$

is invertible and hence $\mathcal{B}$ and $\mathcal{C}$ must have full rank. In particular this implies $d_b \geq r+1$ and $d_c \geq r+1$. Let

$$\overline{\mathcal{B}} := \left[N^{r+1}b, N^{r+2}b, \ldots, N^{d_b-1}b\right] \in \mathbb{R}^{n\times(d_b-r-1)},$$

then by the definition of d_b the matrix $\left[\overline{\mathcal{B}}, \mathcal{B}\right]$ has full column rank.

Step 1c: The matrix $\mathcal{L}$.
If $d_c = r+1$, then $\mathcal{L}$ is the empty matrix. Otherwise let

$$\overline{\mathcal{C}} := \left[cN^{d_c-1}/cN^{d_c-2}/\cdots/cN^{r+1}\right] \in \mathbb{R}^{(d_c-r-1)\times n}.$$

Then $\ker\left[\overline{\mathcal{C}}/\mathcal{C}\right]$ is an $(n-d_c)$-dimensional subspace of $\ker\mathcal{C}$ (where $\dim\ker\mathcal{C} = n-r-1$), i.e. there exists a full rank matrix $L \in \mathbb{R}^{n\times(d_c-r-1)}$ such that $\operatorname{im} L \oplus \ker\left[\overline{\mathcal{C}}/\mathcal{C}\right] = \ker\mathcal{C}$. In particular $\operatorname{im} L \cap \ker\overline{\mathcal{C}} = \{0\}$ and $\operatorname{im} L \subseteq \ker\mathcal{C}$. Let

$$\mathcal{L} := \gamma L\big(\overline{\mathcal{C}}L\big)^{-1}.$$

It remains to show that, firstly, $\mathcal{L}$ is well defined, i.e. that $\overline{\mathcal{C}}L$ is an invertible matrix, and, secondly, that $[\mathcal{L}, \overline{\mathcal{B}}, \mathcal{B}]$ has full rank (otherwise the matrix S is not well defined). Assume that $\overline{\mathcal{C}}Lm = 0$ for some $m \in \mathbb{R}^n$. Then $Lm \in \operatorname{im} L \cap \ker\overline{\mathcal{C}} = \{0\}$, hence $\overline{\mathcal{C}}L$ has only a trivial kernel which implies invertibility. To show that $[\mathcal{L}, \overline{\mathcal{B}}, \mathcal{B}]$ has full rank, observe that $\operatorname{im}\mathcal{L} = \operatorname{im} L$ and, by the definition of the relative degree, $\operatorname{im}\left[\overline{\mathcal{B}}, \mathcal{B}\right] \subseteq \ker\overline{\mathcal{C}}$. Hence $\{0\} = \operatorname{im} L \cap \ker\overline{\mathcal{C}} \supseteq \operatorname{im}\mathcal{L} \cap \operatorname{im}\left[\overline{\mathcal{B}}, \mathcal{B}\right]$, which implies that $[\mathcal{L}, \overline{\mathcal{B}}, \mathcal{B}]$ has full rank.

Step 1d: The matrix $\overline{\mathcal{L}}$.
If $d_b = r+1$, then $\overline{\mathcal{L}}$ is the empty matrix. Otherwise choose, analogously as in the previous step, a full rank matrix $K \in \mathbb{R}^{(d_b-r-1)\times n}$ such that $\operatorname{im} K^\top \oplus \ker\left[\overline{\mathcal{B}}, \mathcal{B}\right]^\top = \ker\mathcal{B}^\top$. Again the matrix $\overline{\mathcal{B}}^\top K^\top$ is invertible. Let

$$\mathcal{K} = (K\overline{\mathcal{B}})^{-1}K,$$

with an analogous argument as in Step 1c it can be shown that $\left[\mathcal{K}/\overline{\mathcal{C}}/\mathcal{C}\right]$ has full rank, hence it is possible to choose a full rank matrix $\overline{\mathcal{L}} \in \mathbb{R}^{n\times(n-d_c-d_b+r+1)}$ such that

$$\operatorname{im}\overline{\mathcal{L}} = \ker\left[\mathcal{K}/\overline{\mathcal{C}}/\mathcal{C}\right].$$

It remains to show that $\left[\mathcal{L}, \overline{\mathcal{L}}, \overline{\mathcal{B}}, \mathcal{B}\right]$ has full rank (i.e. is invertible). To show this, first observe that, by the definition of the relative degree, $\operatorname{im}\mathcal{B} \cap \ker\mathcal{C} = \{0\}$ and recall that $\operatorname{im}\mathcal{L} \cap \ker\overline{\mathcal{C}} = \{0\}$ and analogously $\operatorname{im}\mathcal{K}^\top \cap \ker\overline{\mathcal{B}}^\top = \{0\}$, the latter is equivalent to $\operatorname{im}\overline{\mathcal{B}} \cap \ker\mathcal{K} = \{0\}$. Altogether this yields

$$\ker\left[\mathcal{K}/\overline{\mathcal{C}}/\mathcal{C}\right] \cap \operatorname{im}\left[\mathcal{L}, \overline{\mathcal{B}}, \mathcal{B}\right] = \{0\},$$

which implies that the square matrix $\left[\mathcal{L}, \overline{\mathcal{L}}, \overline{\mathcal{B}}, \mathcal{B}\right]$ has full rank which completes the first step of the proof.

Step 2: The normal form is obtained by the transformation matrices S and T.
It will now be shown that the products ST, Sb, cT and SNT have the desired form.

Step 2a: $ST = \widehat{I}$.
By definition $ST = \widehat{I}$.

Step 2b: $Sb = \widehat{b}$.
Let $e_r = [0, \ldots, 0, 1, \underbrace{0, \ldots, 0}_{r}]^\top \in \mathbb{R}^n$, then $Sb = \widehat{b} = \gamma e_r$ if, and only if, $b = \gamma S^{-1} e_r$. The latter is fulfilled since

$$\gamma S^{-1} = \left[\mathcal{L}, \overline{\mathcal{L}}, \overline{\mathcal{B}}, \underbrace{b, Nb, \ldots, N^r b}_{=\mathcal{B}}\right].$$

Step 2c: $cT = \widehat{c}$.
Choose a full rank matrix $\overline{\mathcal{K}} \in \mathbb{R}^{(n-d_c-d_b+r+1)\times n}$ such that

$$\operatorname{im}\overline{\mathcal{K}}^\top = \ker\left[\mathcal{L}, \overline{\mathcal{B}}, \mathcal{B}\right]^\top.$$

It can be shown analogously as in Step 1d that the square matrix $\mathfrak{C} := \left[\overline{\mathcal{C}}/\overline{\mathcal{K}}/\mathcal{K}/\mathcal{C}\right]$ has full rank (i.e. is invertible). Writing $\mathfrak{B} := \left[\mathcal{L}, \overline{\mathcal{L}}, \overline{\mathcal{B}}, \mathcal{B}\right]$ the matrix T can clearly be written as

$$T = \mathfrak{C}^{-1}\gamma^{-1}\mathfrak{C}\mathfrak{B}\widehat{I}.$$

Since $c\mathfrak{C}^{-1} = [0, \ldots, 0, 1]$ it remains to show that

$$[0, \ldots, 0, 1]\gamma^{-1}\mathfrak{C}\mathfrak{B}\widehat{I} = [0, \ldots, 0, 1] = \widehat{c},$$

or, equivalently, that the last row of the product $\mathfrak{C}\mathfrak{B}$ equals the last row of $\gamma\widehat{I}^{-1}$. It is easy to see that the last row of $\gamma\widehat{I}^{-1}$ is

$$[0, \ldots, 0, cb, cNb, \ldots, cN^r b].$$

Observe that

$$\mathfrak{C}\mathfrak{B} = \begin{bmatrix} \overline{\mathcal{C}}\mathcal{L} & \overline{\mathcal{C}}\overline{\mathcal{L}} & \overline{\mathcal{C}}\overline{\mathcal{B}} & \overline{\mathcal{C}}\mathcal{B} \\ \overline{\mathcal{K}}\mathcal{L} & \overline{\mathcal{K}}\overline{\mathcal{L}} & \overline{\mathcal{K}}\overline{\mathcal{B}} & \overline{\mathcal{K}}\mathcal{B} \\ \mathcal{K}\mathcal{L} & \mathcal{K}\overline{\mathcal{L}} & \mathcal{K}\overline{\mathcal{B}} & \mathcal{K}\mathcal{B} \\ \mathcal{C}\mathcal{L} & \mathcal{C}\overline{\mathcal{L}} & \mathcal{C}\overline{\mathcal{B}} & \mathcal{C}\mathcal{B} \end{bmatrix}. \tag{5.3.4}$$

The matrices $\mathcal{L}$ and $\overline{\mathcal{L}}$ are such that $\operatorname{im}\mathcal{L}$ and $\operatorname{im}\overline{\mathcal{L}}$ are both subspaces of $\ker\mathcal{C}$, hence $\mathcal{C}\mathcal{L} = 0$ and $\mathcal{C}\overline{\mathcal{L}} = 0$. From the definition of the relative degree it follows that $\mathcal{C}\overline{\mathcal{B}} = 0$. Together with (5.3.3) this shows that the last row of $\mathfrak{C}\mathfrak{B}$ is $[0, \ldots, 0, cb, cNb, \ldots, cN^r b]$.

Step 2d: $SNT = \widehat{N}$.
Invoking the notation of Step 2c write

$$SNT = (\mathfrak{C}\mathfrak{B})^{-1}\mathfrak{C}N\mathfrak{B}\widehat{I}.$$

Note that the product $\mathfrak{C}\mathfrak{B}$ in (5.3.4) can further be simplified by the following observations, $\overline{\mathcal{C}}\mathcal{L} = \gamma I$, $\overline{\mathcal{C}}[\overline{\mathcal{L}}, \overline{\mathcal{B}}, \mathcal{B}] = 0$, $\overline{\mathcal{K}}[\mathcal{L}, \overline{\mathcal{B}}, \mathcal{B}] = 0$, $\mathcal{K}[\overline{\mathcal{L}}, \mathcal{B}] = 0$, and $\mathcal{K}\overline{\mathcal{B}} = I$:

$$\mathfrak{C}\mathfrak{B} = \begin{bmatrix} \gamma I & 0 & 0 & 0 \\ 0 & \overline{\mathcal{K}}\overline{\mathcal{L}} & 0 & 0 \\ \mathcal{K}\mathcal{L} & 0 & I & 0 \\ 0 & 0 & 0 & \mathcal{C}\mathcal{B} \end{bmatrix}.$$

Hence

$$(\mathfrak{C}\mathfrak{B})^{-1} = \begin{bmatrix} \gamma^{-1} I & 0 & 0 & 0 \\ 0 & (\overline{\mathcal{K}}\overline{\mathcal{L}})^{-1} & 0 & 0 \\ -\gamma^{-1}\mathcal{K}\mathcal{L} & 0 & I & 0 \\ 0 & 0 & 0 & (\mathcal{C}\mathcal{B})^{-1} \end{bmatrix}.$$

By Proposition 5.3.7,

$$N\overline{\mathcal{B}} = \overline{\mathcal{B}} \left[\begin{smallmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{smallmatrix}\right] \quad \text{and} \quad \overline{\mathcal{C}}N = \left[\begin{smallmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{smallmatrix}\right] \overline{\mathcal{C}},$$

furthermore $\mathcal{C}N\mathcal{B} = \mathcal{C}\mathcal{B}\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right]$, $\overline{\mathcal{K}}N\mathcal{B} = 0$ and $\mathcal{C}N\overline{\mathcal{L}} = 0$, hence

$$\mathfrak{C}N\mathfrak{B} = \begin{bmatrix} \gamma\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] & 0 & 0 & 0 \\ \overline{\mathcal{K}}N\mathcal{L} & \overline{\mathcal{K}}N\overline{\mathcal{L}} & 0 & 0 \\ \mathcal{K}N\mathcal{L} & \mathcal{K}N\overline{\mathcal{L}} & \left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] & \mathcal{K}N\mathcal{B} \\ \mathcal{C}N\mathcal{L} & 0 & 0 & \mathcal{C}\mathcal{B}\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] \end{bmatrix}.$$

Therefore,

$$SNT = (\mathfrak{C}\mathfrak{B})^{-1}\mathfrak{C}N\mathfrak{B}\widehat{I} = \begin{bmatrix} \left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] & 0 & 0 & 0 \\ \widehat{E_1} & \widehat{N_1} & 0 & 0 \\ E_2 & \widehat{E_3} & \left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] & \mathcal{K}N\mathcal{B}I_* \\ (\mathcal{C}\mathcal{B})^{-1}\mathcal{C}N\mathcal{L} & 0 & 0 & \left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] I_* \end{bmatrix},$$

where $\widehat{E_1} = (\overline{\mathcal{K}}\mathcal{L})^{-1}\overline{\mathcal{K}}N\mathcal{L}$, $\widehat{N_1} = (\overline{\mathcal{K}}\mathcal{L})^{-1}\overline{\mathcal{K}}N\overline{\mathcal{L}}$, $E_2 = -\mathcal{K}\mathcal{L}\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right] + \mathcal{K}N\mathcal{L}$, $\widehat{E_3} = \mathcal{K}N\overline{\mathcal{L}}$, and I_* is given by (5.3.2). Note that

$$\mathcal{C}N\mathcal{L} = \begin{bmatrix} cN^{r+1} \\ cN^r \\ \vdots \\ cN^2 \\ cN \end{bmatrix}\mathcal{L} = \begin{bmatrix} \gamma cN^{r+1}L(\overline{\mathcal{C}}L)^{-1} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = \gamma\begin{bmatrix} 0 & \cdots & 0 & 1 \\ & & 0 & \end{bmatrix},$$

hence $(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}N\mathcal{L} = 0_*$, and

$$\begin{aligned} \mathcal{K}N\mathcal{B} &= \mathcal{K}[Nb, N^2b, \ldots, N^rb, N^{r+1}b] \\ &= [0, 0, \ldots, 0, (K\overline{\mathcal{B}})^{-1}KN^{r+1}b] = \begin{bmatrix} & 1 \\ & 0 \\ 0 & \vdots \\ & 0 \end{bmatrix}, \end{aligned}$$

hence $\mathcal{K}N\mathcal{B}I_* = 0^*$.

Clearly, $\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix} I_* = \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix}$ and it remains to show that $\widehat{N_1}$ is nilpotent. This follows from the fact that $SNT\widehat{I}^{-1} = SNS^{-1}$ is nilpotent and because of the special block structure this implies that $\widehat{N_1}$ must also be nilpotent. Without changing the block structure it is possible to transform $\widehat{N_1}$ to Jordan form N_1, this changes $\widehat{E_1}$ and $\widehat{E_3}$ to E_1 and E_3. $\boxed{\text{qed}}$

Remark 5.3.11 (Explicit calculation of normal form)
The proof of Theorem 5.3.10 is constructive. In fact, for a given pure DAE in standard form (c, N, I, b) with negative relative degree $r \geq 0$ and impulse-controllability- and impulse-observability-indices $d_b, d_c \in \mathbb{N}$, respectively, the specific matrices in the normal form are given as follows:

$$\begin{aligned}
E_1 &= J^{-1}(\overline{\mathcal{K}\mathcal{L}})^{-1}\overline{\mathcal{K}}N\mathcal{L} \in \mathbb{R}^{(n-d_c-d_b+r+1)\times(d_b-r-1)},\\
E_2 &= \mathcal{K}N\mathcal{L} - \mathcal{K}\mathcal{L}\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix} \in \mathbb{R}^{(d_b-r-1)\times(d_c-r-1)},\\
E_3 &= \mathcal{K}N\overline{\mathcal{L}}J \in \mathbb{R}^{(d_b-r-1)\times(n-d_c-d_b+r+1)},\\
N_1 &= J^{-1}(\overline{\mathcal{K}\mathcal{L}})^{-1}\overline{\mathcal{K}}N\overline{\mathcal{L}}J \in \mathbb{R}^{(n-d_c-d_b+r+1)\times(n-d_c-d_b+r+1)},\\
0^* &= \begin{bmatrix} \frac{cb}{-\gamma} & \frac{cNb}{-\gamma} & \cdots & \frac{cN^{r-1}b}{-\gamma} & 1 \\ & & & & \\ & & 0 & & \\ & & & & \end{bmatrix} \in \mathbb{R}^{(d_b-r-1)\times(r+1)},\\
0_* &= (\mathcal{C}\mathcal{B})^{-1}\gamma \begin{bmatrix} & 1 \\ & 0 \\ 0 & \vdots \\ & 0 \end{bmatrix} \in \mathbb{R}^{(r+1)\times(d_c-r-1)},\\
I_* &= \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & & \ddots & \\ \frac{cb}{-\gamma} & \frac{cNb}{-\gamma} & \cdots & \frac{cN^{r-1}b}{-\gamma} & 1 \end{bmatrix} \in \mathbb{R}^{(r+1)\times(r+1)},
\end{aligned}$$

where

$$
\begin{aligned}
\mathcal{B} &:= \left[b, Nb, \ldots, N^r b\right], \quad \overline{\mathcal{B}} := \left[N^{r+1}b, N^{r+2}b, \ldots, N^{d_b-1}b\right] \\
\overline{\mathcal{C}} &:= \left[cN^{d_c-1}/cN^{d_c-2}/\cdots/cN^{r+1}\right], \quad \mathcal{C} := \left[cN^r/\ldots/cN/c\right], \\
\mathcal{K} &:= [0, I]\left(\left[\mathcal{B}^\top/\overline{\mathcal{B}}^\top\right][\mathcal{B}, \overline{\mathcal{B}}]\right)^{-1}\left[\mathcal{B}^\top/\overline{\mathcal{B}}^\top\right] \in \mathbb{R}^{(d_b-r-1)\times n} \\
\mathcal{L} &:= \gamma\left[\overline{\mathcal{C}}^\top, \mathcal{C}^\top\right]\left([\overline{\mathcal{C}}/\mathcal{C}]\left[\overline{\mathcal{C}}^\top, \mathcal{C}^\top\right]\right)^{-1}[I/0] \in \mathbb{R}^{n\times(d_c-r-1)}, \\
\overline{\mathcal{K}}^\top &\in \mathbb{R}^{n\times(n-d_c-d_b+r+1)} \text{ is a basis of } \ker\left[\mathcal{L}^\top/\overline{\mathcal{B}}^\top/\mathcal{B}^\top\right], \\
\overline{\mathcal{L}} &\in \mathbb{R}^{n\times(n-d_c-d_b+r+1)} \text{ is a basis of } \ker\left[\mathcal{K}/\overline{\mathcal{C}}/\mathcal{C}\right],
\end{aligned}
$$

and $J \in \mathbb{R}^{(n-d_c-d_b+r+1)\times(n-d_c-d_b+r+1)}$ is a basis transformation such that N_1 is in Jordan normal form.

If the DAE (c, E, A, b) is not in standard form, then either N and b in the above formulae must be replaced by $A^{-1}E$ and $A^{-1}b$, resp., or N and c must be replaced by EA^{-1} and cA^{-1}, resp. □

The normal form of Theorem 5.3.10 can be viewed as a specialization of (2-5.4) in [Dai89, p. 52]: it is more explicit and simpler, the size of the different blocks is explicitly given, and the influence of the input on the states can be seen more directly as well as the influence of the states on the output. Furthermore no proof is given in [Dai89].

With the normal form from Theorem 5.3.10 it is now possible to give characterization of impulse-controllability and -observability.

Theorem 5.3.12 (Impulse-controllability and -observability) Consider a pure DAE (c, E, A, b) with negative relative degree $r \geq 0$ and impulse-controllability- and impulse-observability-indices $d_b, d_c \in \mathbb{N}$.

Let $N_1 \in \mathbb{R}^{(n-d_b-d_c+r+1)\times(n-d_b-d_c+r+1)}$ be given as in Theorem 5.3.10. Then the following characterizations of impulse-controllability and -observability hold:

(i) The DAE is impulse-controllable if, and only if, $d_c = r+1$ and $N_1 = 0$

(ii) The DAE is impulse-observable if, and only if, $d_b = r+1$ and $N_1 = 0$. □

Proof. Let $(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b})$ be the normal form of (c, E, A, b) from Theorem 5.3.10.

(i) It is easily seen that $(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b})$ is equivalent to $(\widetilde{c}, \widetilde{N}, I, \widetilde{b})$ with

$$\widetilde{N} = \begin{bmatrix} \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} & 0 & 0 & 0 \\ E_1 & N_1 & 0 & 0 \\ 0_* & 0 & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} & 0 \\ E_2 & E_3 & \widetilde{0^*} & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} \end{bmatrix}$$

where the matrices $E_1, E_2, E_3, N_1, 0_*$ are the same as in Theorem 5.3.10 and $\widetilde{0^*}$ has the same structure as 0^* from Theorem 5.3.10, in particular 0_* and $\widetilde{0^*}$ have a one in the upper right corner. The vector $\widetilde{b}$ is given by $\widetilde{b} = [0, \ldots, 0, \gamma, 0, \ldots, 0]^\top$ with $\gamma \neq 0$ at the $(n - d_b + 1)$-th position.

By Proposition 5.1.4 the DAE $(\widetilde{c}, \widetilde{N}, I, \widetilde{b})$ is impulse-controllable if, and only if, $\operatorname{im}[\widetilde{N}\widetilde{b}, \widetilde{N}^2\widetilde{b}, \ldots, \widetilde{N}^{n-1}\widetilde{b}] = \operatorname{im}\widetilde{N}$. It is easily seen that

$$\operatorname{im}[\widetilde{N}\widetilde{b}, \widetilde{N}^2\widetilde{b}, \ldots, \widetilde{N}^{n-1}\widetilde{b}] = \operatorname{im} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} & 0 \\ \widetilde{0^*} & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} \end{bmatrix}.$$

hence for the given DAE impulse-controllability is equivalent to the condition

$$\operatorname{im} \begin{bmatrix} \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} & 0 \\ E_1 & N_1 \\ 0_* & 0 \\ E_2 & E_3 \end{bmatrix} \subseteq \operatorname{im} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} & 0 \\ \widetilde{0^*} & \begin{smallmatrix}0&&&\\1&\ddots&&\\&\ddots&\ddots&\\&&1&0\end{smallmatrix} \end{bmatrix}.$$

A necessary and sufficient condition for this is that the matrix 0_* is not existent (because it has a one in the upper right corner), i.e. $d_c = r + 1$, and that N_1 is the zero matrix.

(ii) It is easily seen that $(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{c})$ is equivalent to $(\widetilde{c}, \widetilde{N}, I, \widetilde{b})$ with

$$\widetilde{N} = \begin{bmatrix} \begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix} & 0 & 0 & 0 \\ \widetilde{0_*} & \widetilde{\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right]} & 0 & 0 \\ E_1 & 0 & N_1 & 0 \\ E_2 & 0^* & E_3 & \begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix} \end{bmatrix},$$

where E_1, E_2, E_3, N_1, O^* are as in the normal form in Theorem 5.3.10, $\widetilde{0_*}$ has the same structure as 0_* from Theorem 5.3.10 and

$$\widetilde{\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right]} = \begin{bmatrix} 0 & & & & \\ 1 & 0 & & & \\ & \ddots & \ddots & & \\ & & 1 & 0 & \\ * & \cdots & * & 1 & 0 \end{bmatrix}.$$

The vector $\widetilde{c}$ is given by $\widetilde{c} = [0, \ldots, 0, 1, 0, \ldots, 0]$ where the one is at position d_c.

By Proposition 5.2.5 the DAE $(\widetilde{c}, \widetilde{N}, I, \widetilde{b})$ is impulse-observable if, and only if, $\ker[\widetilde{c}\widetilde{N}^{d_c}/\widetilde{c}\widetilde{N}^{d_c-1}/\ldots/\widetilde{c}\widetilde{N}] = \ker \widetilde{N}$. Easy calculations show that (here it is needed that $\widetilde{0_*}$ has a one in the upper right corner)

$$\ker[\widetilde{c}\widetilde{N}^{d_c}/\widetilde{c}\widetilde{N}^{d_c-1}/\ldots/\widetilde{c}\widetilde{N}] = \ker \begin{bmatrix} \begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix} & 0 & 0 & 0 \\ \widetilde{0_*} & \widetilde{\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right]} & 0 & 0 \end{bmatrix}.$$

Hence impulse-controllability of the DAE is equivalent to the condition

$$\ker \begin{bmatrix} \begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix} & 0 & 0 & 0 \\ \widetilde{0_*} & \widetilde{\left[\begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix}\right]} & 0 & 0 \end{bmatrix} \subseteq \ker \begin{bmatrix} E_1 & 0 & N_1 & 0 \\ E_2 & 0^* & E_3 & \begin{smallmatrix}0&&\\1&\ddots&\\&1&0\end{smallmatrix} \end{bmatrix}.$$

Because 0^* has a one in the upper right corner the inclusion holds if, and only if, 0^* does not exists, i.e. $d_b = r + 1$, and $N_1 = 0$. $\boxed{\text{qed}}$

Corollary 5.3.13 (Special normal form)
The pure DAE (c, E, A, b) is impulse-controllable and -observable if, and only if, the normal form $(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b})$ from Theorem 5.3.10 reduces to

$$(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b}) = \left([0, \ldots, 0, 1], \begin{bmatrix} 0 & 0 \\ 0 & \begin{smallmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{smallmatrix} \end{bmatrix}, \begin{bmatrix} I & 0 \\ 0 & \begin{smallmatrix} 1 & & \\ & \ddots & \\ * & \cdots * & 1 \end{smallmatrix} \end{bmatrix}, [\underbrace{0, \ldots, 0}_{n-r-1}, \gamma, \underbrace{0, \ldots, 0}_{r}]^\top \right).$$

In particular, if the negative relative degree r of a pure DAE (c, E, A, b) is maximal, i.e. $r = n - 1$, then (c, E, A, b) is impulse-controllable and -observable. □

Remark 5.3.14 ((Un-)controllable and (un-)observable substates)
For a pure DAE in normal form $(\hat{c}, \hat{N}, \hat{I}, \hat{b})$ as in Theorem 5.3.10 divide the corresponding state variable x into $x = (x_1/x_2/x_3/x_4)$ conforming to the block sizes of the normal form. First observe that the substates x_1 and x_2 are independently given from the input u, in fact, x_1 is the unique solution of the regular DAE (or the corresponding ITP)

$$\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix} \dot{x}_1 = x_1$$

and w is the unique solution of a regular DAE with inhomogeneity $E_1 x_1$:

$$N_1 \dot{x}_2 = x_2 - E_1 x_1.$$

Therefore, x_1 and x_2 can be called *uncontrollable substates.* The state x_4 is uniquely given by the regular DAE

$$\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix} \dot{x}_4 = I_* x_4 + [\gamma, 0, \ldots, 0]^\top u - 0_* x_1,$$

and finally the state x_3 is the unique solution of the regular DAE

$$\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix} \dot{x}_3 = x_3 - E_2 x_1 - E_3 x_2 - 0^* x_4.$$

Therefore, the state x_4 can be directly controlled by u and the state x_3 can be controlled by u via the state ω, hence the states x_3 and x_4

can be called *(impulse-) controllable substates*. In fact, if the states x_1 and x_2 do not exists, then Proposition 5.3.12 shows that the pure DAE is impulse-controllable.

The output y is given by $y = [0, \ldots, 0, 1]x_4$ and since x_4 does not depend on the states x_2 and x_3, the latter can be called *unobservable substates*. If the states x_2 and x_3 do not exist, then, by Theorem 5.3.12, the DAE is impulse-observable, hence the states x_1 and x_4 can be called *(impulse-)observable substates*. This relation between the different substates is also illustrated in Figure 8. □

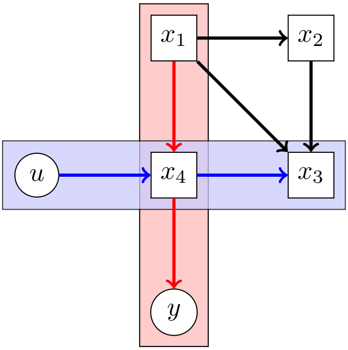

Figure 8: Dependencies of the four substates corresponding to the normal form from Theorem 5.3.10, see also Remark 5.3.14. The substate x_1 is (impulse-) observable but not controllable, x_2 is neither controllable nor observable, x_3 is (impulse-) observable but not observable and x_4 is both, impulse-controllable and -observable.

With the help of the normal form from Theorem 5.3.10 it is possible to explicitly define an input u such that, for the impulse-controllable case, each ITP has a impulse free solution. For the impulse-observable case it is possible to explicitly calculate the impulses in the states if the output is known.

Theorem 5.3.15 (Impulse-elimination and -reconstruction)
Consider a pure DAE in standard form (c, N, I, b) and with the normal form $(\widehat{c}, \widehat{N}, \widehat{I}, \widehat{b})$ as in Theorem 5.3.10.

(i) Assume (c, N, I, b) is impulse-controllable, i.e. the DAE is equivalent to the *impulse-controllable normal form*

$$\begin{bmatrix} 0 & 0 \\ 0 & \begin{smallmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix} \\ E_3 & \end{bmatrix} \dot{z} = z + \begin{bmatrix} 0 \\ cN^r b \\ 0 \\ \vdots \\ 0 \end{bmatrix} u,$$

$$y = [0, \ldots, 0, \tfrac{cb}{cN^r b}, \tfrac{cNb}{cN^r b}, \ldots, \tfrac{cN^{r\text{-}1}b}{cN^r b}, 1, 0, \ldots, 0]$$

where the size of the submatrix $\left[\begin{smallmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix}\right]$ is $d_b \times d_b$. For some initial trajectory $z^0 = [z_1^0 / z_2^0] \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{(n-d_b)+d_b}$ and initial time $t_0 \in \mathbb{R}$ let

$$\Delta z := -z_2^0(t_0-) - \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}^\top \begin{bmatrix} 0 \\ E_3 \end{bmatrix} z_1^0(t_0-) \in \mathbb{R}^{d_b},$$

then the input $u = u_{\mathbb{D}}^{\mathrm{reg}}$ with

$$u^{\mathrm{reg}}(t) := \begin{cases} 0 & , \quad t < t_0, \\ \frac{1}{cN^r b}[1, t - t_0, \frac{(t-t_0)^2}{2}, \ldots, \frac{(t-t_0)^{(d_b-1)}}{(d_b-1)!}]\Delta z & , \quad t \geq t_0 \end{cases}$$

ensures that the unique solution z of the corresponding ITP is impulse free at t_0.

(ii) Assume (c, N, I, b) is impulse-observable, i.e. the DAE is equivalent to the *impulse-observable normal form*

$$\begin{bmatrix} 0 & E_1 & 0 \\ 0 & \multicolumn{2}{c}{\begin{smallmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix}} \end{bmatrix} \dot{z} = z + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ cN^r b \\ cN^{r\text{-}1} b \\ \vdots \\ cb \end{bmatrix} u,$$

$$y = [0, 0, \ldots, 0, 1]x,$$

where the size of the submatrix $\left[\begin{smallmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{smallmatrix}\right]$ is $d_c \times d_c$. For some initial time $t_0 \in \mathbb{R}$, assume $u^{(*)}(t_0+) = 0$ and let $z \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^n$

be some ITP solution of the above DAE in impulse-observable normal form. Choose $k \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_k \in \mathbb{R}$ such that

$$y[t_0] + \sum_{i=0}^{r} cN^i bu[t_0]^{(i)} = \sum_{j=0}^{k} \alpha_j \delta_{t_0}^{(j)},$$

and decompose z into $z = [z_1/z_2] \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{(n-dc)+dc}$, then

$$z_2[t_0] = \sum_{i=0}^{d_c-2} \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & \ddots & \ddots \\ & & 1 \;\; 0 \end{bmatrix}^i \begin{bmatrix} 0 \\ \alpha_{d_c-2} \\ \vdots \\ \alpha_1 \\ \alpha_0 \end{bmatrix} \delta_{t_0}^{(i)}.$$

and

$$z_1[t_0] = [E_1, 0] z_2[t_0]'.$$

□

Proof. (i) First note that the product $0^* I_*^{-1}$ is a zero matrix with a single one at the right upper corner and that $\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix} I_*^{-1} = \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix}$. Hence the coordinate transformation $x \mapsto I_* x$ and a permutation of the blocks yields the impulse-controllable normal form. Let $z = [z_1/z_2]$ be the unique solution of the ITP corresponding to the DAE in impulse-controllable normal form with initial trajectory z^0 and input u. Then, clearly, $z_1 = z^0_{1\,(-\infty,t_0)}$, in particular $z_1[t_0] = 0$. The solution formula from Corollary 3.4.4 implies

$$z_2 = -\sum_{i=0}^{d_b-1} \left(\begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix}_{[t_0,\infty)} \frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} \right)^i \left(-z^0_{2\,(-\infty,t_0)} + \begin{bmatrix} cN^r b \\ 0 \\ \vdots \\ 0 \end{bmatrix} u \right.$$
$$\left. - \begin{bmatrix} 0 \\ E_3 \end{bmatrix} z_1^0(t_0-)\delta_{t_0} \right),$$

hence

$$z_2[t_0] = -\begin{bmatrix} 0 \\ cN^r bu'[t_0] \\ cN^r bu''[t_0] \\ \vdots \\ cN^r bu^{(d_b-1)}[t_0] \end{bmatrix} - \sum_{i=0}^{d_b-1} \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix}^j \begin{bmatrix} 0 \\ E_3 \end{bmatrix} z_1^0(0-)\delta_0^{(i)}$$
$$- \sum_{i=1}^{d_b-1} \begin{bmatrix} 0 & & \\ 1 & \ddots & \\ & 1 & 0 \end{bmatrix}^i z_2^0(0-)\delta_0^{(i-1)}.$$

By construction, for $i = 1, \ldots, d_b - 1$,

$$\begin{aligned} cN^r bu^{(i)}[t_0] &= -\sum_{j=1}^{i} [\underbrace{0,\ldots,0}_{j-1},1,0,\ldots,0]\Delta z\ \delta_{t_0}^{(i-j)} \\ &= -\sum_{j=1}^{i} [\underbrace{0,\ldots,0}_{j},1,0,\ldots,0] \begin{bmatrix} 0 \\ E_3 \end{bmatrix} z_1^0(t_0-)\delta_{t_0}^{(i-j)} \\ &\quad -\sum_{j=1}^{i} [\underbrace{0,\ldots,0}_{j-1},1,0,\ldots,0] z_2^0(t_0-)\delta_{t_0}^{(i-j)} \end{aligned}$$

and with the observation that the zero matrix in $\begin{bmatrix} 0 \\ E_3 \end{bmatrix}$ cannot vanish (unless E_1 does not exist at all) it follows that each "row" of $z_2[t_0]$ is zero.

(ii) With the notation of Remark 5.3.11 it follows that $\frac{1}{\gamma}\mathcal{CB}0_*$ is a zero matrix with a single one at the right upper corner and its easy to see that $\frac{1}{\gamma}\mathcal{CB}\left[\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&\ddots\\&&1\;0\end{smallmatrix}\right]\big(\frac{1}{\gamma}\mathcal{CB}I_*\big)^{-1} = \left[\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&\ddots\\&&1\;0\end{smallmatrix}\right]$ and $[0,\ldots,0,1]\big(\frac{1}{\gamma}\mathcal{CB}I_*\big)^{-1} = [0,\ldots,0,1]$, hence together with a permutation of the blocks it is already shown that any impulse-observable pure DAE can be put into the impulse-observable normal form. Let y_0 be the unique output of the corresponding ITP with zero initial trajectory and input u with $u^{(*)}(t_0+) = 0$, then it follows by the solution formula from Corollary 3.4.4 that $y_0[t_0] = -\sum_{i=0}^{r} cN^i bu[t_0]^{(i)}$, hence, by linearity, the term $y[t_0] + \sum_{i=0}^{r} cN^i bu[t_0]^{(0)}$ is just the impulsive part of the output of the pure DAE with an input signal u with $u[t_0] = 0$. Therefore, it suffices to consider $u[t_0] = 0$ in the following. In this case the unique ITP solution $z = [z_1/z_2] \in (\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{(n-dc)+d_c}$ of the pure DAE with some initial trajectory $[z_1^0/z_2^0]$ fulfills

$$\begin{aligned} z_2[t_0] &= -\sum_{i=1}^{d_c-1} \left[\begin{smallmatrix}0&&\\1&\ddots&\\&\ddots&\ddots\\&&1\;0\end{smallmatrix}\right]^i z_2^0(t_0-)\delta_{t_0}^{(i-1)} \\ z_1[t_0] &= [E_1, 0] z_2[t_0]', \end{aligned}$$

Since

$$\begin{aligned}\sum_{i=0}^{k} \alpha_k \delta^{(i)} &= y[t_0] = [0, \dots, 0, 1] z_2[t_0] \\ &= - \sum_{i=0}^{d_c - 2} [0, \dots, 0, 1, \underbrace{0, \dots, 0}_{i+1}] z_2^0(t_0-) \delta_{t_0}^{(i)}\end{aligned}$$

it follows by Remark 2.1.14 that $\alpha_{d_c - 1} = \ldots = \alpha_k = 0$ and

$$\begin{bmatrix} 0 \\ \alpha_{d_c-2} \\ \vdots \\ \alpha_1 \\ \alpha_0 \end{bmatrix} = - \begin{bmatrix} 0 & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix} z_2^0(t_0-),$$

hence the claim is shown. qed

References

[AE05] H. Amann and J. Escher (2005): *Analysis I.* Birkhäuser, Basel Boston Berlin, 1st english ed. Cited on page 98.

[Apl91] J. D. Aplevich (1991): *Implicit Linear Systems.* No. 152 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin. Cited on page 13.

[Ber08] T. Berger (2008): *Zur asymptotischen Stabilität linearer differential-algebraischer Gleichungen.* Bachelor Thesis, Institute for Mathematics, Ilmenau University of Technology. Cited on pages 125 and 137.

[BIT09] T. Berger, A. Ilchmann and S. Trenn (2009): The Quasi-Weierstraß form for regular matrix pencils. *submitted for publication.* Preprint available online, Institute for Mathematics, Ilmenau University of Technology, Preprint Number 09-21. Cited on page 18.

[BM02] K. Balla and R. März (2002): A unified approach to linear differential algebraic equations and their adjoints. *Z. Anal. Anwend.*, **21**, pp. 783–802. Cited on page 13.

[Cam80] S. L. Campbell (1980): *Singular Systems of Differential Equations I.* Pitman, New York. Cited on page 13.

[Cam82] S. L. Campbell (1982): *Singular Systems of Differential Equations II.* Pitman, New York. Cited on page 13.

[Cam87] S. L. Campbell (1987): A general form for solvable linear time varying singular systems of differential equations. *SIAM J. Math. Anal.*, **18**(4), pp. 1101–1115. doi: 10.1137/0518081. Cited on page 70.

[Cob84] J. D. Cobb (1984): Controllability, observability and duality in singular systems. *IEEE Trans. Autom. Control*, **AC-29**, pp. 1076–1082. Cited on pages 14, 15, 144, 147, 150, and 152.

[CP83] S. L. Campbell and L. R. Petzold (1983): Canonical forms and solvable singular systems of differential equations. *SIAM J. Alg. & Disc. Meth.*, **4**, pp. 517–521. Cited on pages 17, 56, and 98.

[Dai89] L. Dai (1989): *Singular Control Systems.* No. 118 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Berlin. Cited on pages 13, 17, 141, 144, 147, 150, 152, and 164.

[Dol64] V. Doležal (1964): The existence of a continuous basis of a certain linear subspace of E_r, which depends on a parameter. *Časopis pro pestovani matematiky*, **89**, pp. 466–468. Cited on page 63.

[FLZ92] B. Fisher and C. Lin-Zhi (1992): The product of distributions on R^n. *Commentat. Math. Univ. Carol.*, **33**(4), pp. 605–614. Cited on page 48.

[Fuc68] B. Fuchssteiner (1968): Eine assoziative Algebra über einen Unterraum der Distributionen. *Mathematische Annalen*, **178**, pp. 302–314. Cited on pages 16 and 43.

[Fuc84] B. Fuchssteiner (1984): Algebraic foundation of some distribution algebras. *Studia Mathematica*, **76**, pp. 439–453. Cited on pages 16 and 43.

[Gan59] F. R. Gantmacher (1959): *The Theory of Matrices (Vol. I & II).* Chelsea, New York. Cited on page 15.

[GM86] E. Griepentrog and R. März (1986): *Differential-algebraic equations and their numerical treatment.* No. 88 in Teubner-Texte zur Mathematik. Teubner, Leipzig. Cited on page 13.

[HS83] M. L. J. Hautus and L. M. Silverman (1983): System structure and singular control. *Lin. Alg. Appl.*, **50**, pp. 369–402. Cited on page 14.

[IM05] A. Ilchmann and V. Mehrmann (2005): A behavioural approach to time-varying linear systems, Part 1: General theory. *SIAM J. Control Optim.*, **44**(5), pp. 1725–1747. Cited on page 74.

[IRT07] A. Ilchmann, E. P. Ryan and P. Townsend (2007): Tracking with prescribed transient behavior for nonlinear systems of known relative degree. *SIAM J. Control Optim.*, **46**(1), pp. 210–230. doi:10.1137/050641946. Cited on page 152.

[Isi95] A. Isidori (1995): *Nonlinear Control Systems.* Communications and Control Engineering Series. Springer-Verlag, Berlin, 3rd ed. Cited on page 152.

[Jan71] L. Jantscher (1971): *Distributionen.* De Gruyter Lehrbuch. Walter de Gruyter, Berlin, New York. Cited on pages 21, 22, 23, 24, 28, and 44.

[Kal60] R. E. Kalman (1960): Contributions on the theory of optimal control. *Bol. Soc. Matem. Mexico*, **II. Ser. 5**, pp. 102–119. Cited on pages 144 and 150.

[Kal62] R. E. Kalman (1962): Canonical structure of linear dynamical systems. *Proc. Nat. Acad. Sci. (USA)*, **48**(4), pp. 596–600. Cited on page 152.

[KM06] P. Kunkel and V. Mehrmann (2006): *Differential-Algebraic Equations. Analysis and Numerical Solution.* EMS Publishing House, Zürich, Switzerland. Cited on pages 13, 70, 97, 114, and 152.

[Kön55] H. König (1955): Multiplikation von Distributionen. I. *Mathematische Annalen*, **128**, pp. 420–452. Cited on page 48.

[KR95] H. König and R. Raeder (1995): *Vorlesung über die Theorie der Distributionen.* No. 6-1 in Annales Universitatis

Saraviensis - Series Mathematicae. Universität des Saarlandes, Saarbrücken. Cited on pages 26 and 27.

[Kro90] L. Kronecker (1890): Algebraische Reduction der Schaaren bilinearer Formen. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften zu Berlin*, pp. 1225–1237. Cited on page 15.

[Lan70] S. Lang (1970): *Linear Algebra.* Addison-Wesley, Reading, Massachusetts, 2nd ed. Fourth printing 1973. Cited on page 155.

[LBS89] V. Lakshmikantham, D. D. Baĭnov and P. S. Simeonov (1989): *Theory of Impulsive Differential Equations.* No. 6 in Series in Modern Applied Mathematics. World Scientific, Singapore, Teaneck, NJ. Cited on page 16.

[Lib03] D. Liberzon (2003): *Switching in Systems and Control.* Systems and Control: Foundations and Applications. Birkhäuser, Boston. Cited on page 102.

[LÖMK91] J. Loiseau, K. Özçaldiran, M. Malabre and N. Karcanias (1991): Feedback canonical forms of singular systems. *Kybernetika*, **27**(4), pp. 289–305. Cited on page 152.

[LT09] D. Liberzon and S. Trenn (2009): On stability of linear switched differential algebraic equations. *to appear in IEEE Conf. Decis. Control 2009, Shanghai, China.* Preprint available online, Institute for Mathematics, Ilmenau University of Technology, Preprint Number 09-04. Cited on page 18.

[Mik66] J. Mikusiński (1966): On the square of the Dirac delta-distribution. *Bull. Acad. Pol. Sci., Sér. sci. math. astr. phys.*, **XIV**(9), pp. 511–513. Cited on page 48.

[OD85] D. H. Owens and D. L. Debeljkovic (1985): Consistency and Liapunov stability of linear descriptor systems: A geometric analysis. *IMA J. Math. Control & Information*, pp. 139–151. Cited on pages 104, 124, and 125.

[Rat97] W. Rath (1997): Derivative and proportional state feedback for linear descriptor systems with variable coefficients. *Lin. Alg. Appl.*, **260**, pp. 273–310. Cited on page 152.

[RR96a] P. J. Rabier and W. C. Rheinboldt (1996): Classical and generalized solutions of time-dependent linear differential-algebraic equations. *Lin. Alg. Appl.*, **245**, pp. 259–293. Cited on pages 17 and 74.

[RR96b] P. J. Rabier and W. C. Rheinboldt (1996): Time-dependent linear DAEs with discontinuous inputs. *Lin. Alg. Appl.*, **247**, pp. 1–29. Cited on pages 14, 61, and 62.

[RR02] P. J. Rabier and W. C. Rheinboldt (2002): Theoretical and numerical analysis of differential-algebraic equations. In: P. G. Ciarlet and J. L. Lions (eds.) *Handbook of Numerical Analysis*, vol. VIII, pp. 183–537. Elsevier Science, Amsterdam, The Netherlands. Cited on page 13.

[Son98] E. D. Sontag (1998): *Mathematical Control Theory, Deterministic Finite Dimensional Systems.* Springer-Verlag, New York, 2nd ed. Cited on page 75.

[Tre08a] S. Trenn (2008): Distributional solution theory for linear DAEs. In: *PAMM - Proc. Appl. Math. Mech.*, vol. 8, pp. 10077–10080. GAMM Annual Meeting 2008, Bremen, Wiley-VCH Verlag GmbH, Weinheim. doi:10.1002/pamm.200810077. Published online: 25 Feb 2009. Cited on page 18.

[Tre08b] S. Trenn (2008): Regularity of distributional differential algebraic equations. *submitted for publication.* Preprint available online, Institute for Mathematics, Ilmenau University of Technology, Preprint Number 08-24. Cited on page 18.

[Tre09a] S. Trenn (2009): Impulse free solutions for switched differential algebraic equations. *submitted for publication.* Preprint available online, Institute for Mathematics, Ilmenau University of Technology, Preprint Number 09-03. Cited on page 18.

[Tre09b] S. Trenn (2009): A normal form for pure differential algebraic systems. *Lin. Alg. Appl.*, **430**(4), pp. 1070 – 1084. doi:10.1016/j.laa.2008.10.004. Cited on page 18.

[VLK81] G. C. Verghese, B. C. Levy and T. Kailath (1981): A generalized state-space for singular systems. *IEEE Trans. Autom. Control*, **AC-26**(4), pp. 811–831. Cited on pages 13 and 152.

[Wal70] H.-F. Walter (1970): Über die Multiplikation von Distributionen in einem Folgenmodell. *Mathematische Annalen*, **189**, pp. 211–221. Cited on page 48.

[Wal94] W. Walter (1994): *Einführung in die Theorie der Distributionen.* BI Wissenschaftsverlag, Mannheim Leipzig Wien Zürich, 3rd ed. Cited on pages 48, 77, and 110.

[Wei68] K. Weierstrass (1868): Zur Theorie der bilinearen und quadratischen Formen. *Monatsh. Akademie. Wiss.*, pp. 310–338. Cited on page 15.

[Wer02] D. Werner (2002): *Funktionalanalysis.* Springer-Verlag, Berlin, 4th ed. Cited on page 21.

[Wil07] J. C. Willems (2007): The behavioral approach to open and interconnected systems. *IEEE Control Systems Magazine*, **27**(6), pp. 46–99. Cited on page 142.

[Won74] K.-T. Wong (1974): The eigenvalue problem $\lambda Tx + Sx$. *J. Diff. Eqns.*, **16**, pp. 270–280. Cited on pages 103 and 104.

[YS81] E. L. Yip and R. F. Sincovec (1981): Solvability, controllability and observability of continuous descriptor systems.

IEEE Trans. Autom. Control, **AC-26**, pp. 702–707. Cited on page 73.

List of symbols and abbreviations

$\mathbb{1}$	$:=\mathbb{1}_{\mathbb{R}}$, the constant unity function
$\mathbb{1}_M$	indicator function of set $M \subseteq \mathbb{R}$, 18
$\oplus$	the direct sum (of two linear subspace)
$\int_M f$	The Lebesgue integral of the measurable function $\mathbb{1}_M f : \mathbb{R} \to \mathbb{R}$ for some measurable $M \subseteq \mathbb{R}$
$\int_{t_0} D$	the unique antiderivative of $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ with $\big(\int_{t_0} D\big)(t_0-) = 0$, 41
$\subseteq, \subset$	subset, proper subset
$\dot{\cup}, \dot{\bigcup}$	disjoint union
$\mathfrak{B}_{\mathrm{imp}}$	impulse-controllability-matrix of a pure DAE, 154
$\mathcal{B}^{t_0}_{(E,A,B)}$	the set of all ITP solutions (x,u) of $E\dot{x} = Ax + Bu$ with initial time $t_0 \in \mathbb{R}$, 142
$\mathcal{B}^{t_0}_{(E,A,B,C,D)}$	the set of all ITP solutions (x,u,y) of $E\dot{x} = Ax + Bu$, $y = Cx + Bu$, with initial time $t_0 \in \mathbb{R}$, 148
$\mathbb{C}$	the complex numbers
$\mathfrak{C}_{(E,A)}$	consistency space corresponding to the regular matrix pair (E,A), 124
$\mathfrak{C}_p$	$:= \mathfrak{C}_{(E_p,A_p)}$, the consistency space of subsystem (E_p,A_p) of the switched DAE (4.1.1), 134
$\mathfrak{C}_{t_0}$	consistency space at t_0 of switched DAE (4.1.1), 133
$\mathfrak{C}_{\mathrm{imp}}$	impulse-observability-matrix of a pure DAE, 154
$\mathcal{C}^\infty$	the space of smooth (arbitrarily often differential) functions $f : \mathbb{R} \to \mathbb{R}$, 21

$\mathcal{C}_0^\infty$	the space of test functions, i.e. smooth functions with bounded support, 21
$\mathcal{C}_{\mathrm{pw}}^\infty$	the space of piecewise-smooth functions, 36
$\mathrm{cl}\, M$	the closure of a set $M \subseteq \mathbb{R}$
$\mathbb{D}$	the space of distributions, i.e. the space of all linear and continuous operators $D : \mathcal{C}_0^\infty \to \mathbb{R}$, 21
$\mathbb{D}_M$	the space of distributions with support contained in $M \subseteq \mathbb{R}$, 24
$\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$	the space of piecewise-smooth distributions, 38
$\mathbb{D}_{\mathrm{pwreg}}$	the space of piecewise-regular distributions, 31
$\mathbb{D}_{\mathrm{reg}}$	the space of regular distributions, 22
$D', \dot{D}$	the derivative of a distribution $D \in \mathbb{D}$, 23
$D^{(n)}$	the n-th derivative of a distribution $D \in \mathbb{D}$, $n \in \mathbb{N}$, 23
$D^{(*)}(t+)$	$:= \big(D(t+), D'(t+), \ldots, D^{(i)}(t+), \ldots\big)$, the sequence of right-sided derivatives of $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ at $t \in \mathbb{R}$, 141
$D_n \to D$	convergence of the sequence $(D_n)_{n\to\infty}$ of distribution with the limit $D \in \mathbb{D}$., 23
D_{reg}	$:= D - D[\cdot]$, the regular part of $D \in \mathbb{D}_{\mathrm{pwreg}}$, 34
D^{reg}	$\in L_{1,\mathrm{loc}}$, the function which induces the regular part D_{reg} of $D \in \mathbb{D}_{\mathrm{pwreg}}$, i.e. $D_{\mathrm{reg}} = (D^{\mathrm{reg}})_{\mathbb{D}}$, 34
$D[t], D[\cdot]$	impulsive part of $D \in \mathbb{D}_{\mathrm{pwreg}}$, 34
$D(t+)$	right sided evaluation of a piecewise-smooth distribution $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ at $t \in \mathbb{R}$, 39
$D(t-)$	left sided evaluation of a piecewise-smooth distribution $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ at $t \in \mathbb{R}$, 39

$\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t}$	the distributional derivative operator $\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} : \mathbb{D} \to \mathbb{D}$, $\frac{\mathrm{d}_{\mathbb{D}}}{\mathrm{d}t} D := D'$, 23
$\frac{\mathrm{d}}{\mathrm{d}t}$	the derivative operator for differentiable functions
DAE	= differential algebraic equation
d_b	impulse-controllability-index of a pure DAE, 154
d_c	impulse-observability-index of a pure DAE, 154
$\delta_t, \delta_t^{(n)}$	(n-th derivative of) the Dirac impulse at $t \in \mathbb{R}$, 25
$\Delta_t\{D\}$	$:= D(t+) - D(t-)$, the jump of a piecewise-smooth distribution $D \in \mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty}$ at $t \in \mathbb{R}$, 39
$f_{\mathbb{D}}$	the regular distribution induced by $f \in L_{1,\mathrm{loc}}$, 22
f_M	$:= \mathbb{1}_M f$, restriction of functions $f : \mathbb{R} \to \mathbb{R}$ to the set $M \subseteq \mathbb{R}$
I	identity matrix in $(\mathbb{D}_{\mathrm{pw}\mathcal{C}^\infty})^{n\times n}$, $(\mathcal{C}^\infty_{\mathrm{pw}})^{n\times n}$ or $\mathbb{R}^{n\times n}$, 52
$\operatorname{im} M$	the image of a matrix $M \in \mathbb{R}^{m\times n}$, 19
ITP	= initial trajectory problem, 55
$\ker M$	kernel of some matrix $M \in \mathbb{R}^{m\times n}$, 19
$L_{1,\mathrm{loc}}$	the space of locally integrable functions, 22
$\mathbb{N}$	$:= \{0, 1, 2, \ldots, \}$, the natural numbers
ODE	= ordinary differential equation, 74
$\Pi_{(E,A)}$	consistency projector of the regular matrix pair (E, A), 109
Π_p	$:= \Pi_{(E_p,A_p)}$, the consistency projector of subsystem (E_p, A_p) of the switched DAE (4.1.1), 111
$\mathbb{R}$	the real numbers

$\mathbb{R}_{>0}, \mathbb{R}_{\geq 0}$	positive, non-negative real numbers
σ	switching signal, 101
$\Sigma^{m\times n}$	system class for distributional DAEs, (3.1.1), 55
$\Sigma_{(t_0,x^0)}$	space of switching signals which guarantee solutions without jumps, 137
Σ^{τ_d}	space of switching signals with dwell time $\tau_d > 0$, 138
$\operatorname{supp} D$	support of $D \in D$, 24
$\operatorname{supp} f$	support of function $f : \mathbb{R} \to \mathbb{R}$, 21
τ_d	dwell time of a switching signal, 138
$\mathcal{V}_i, \mathcal{W}_i$	i-th element of Wong sequences for regular matrix pair (E, A), 103
$\mathcal{V}^*, \mathcal{W}^*$	limit of the Wong sequences of a regular matrix pair (E, A), 105
$\mathbb{Z}$	$:= \{\dots, -2, -1, 0, 1, 2, \dots\}$, the integers

Index